

# Evaluation of Non-Intrusive Load Monitoring for Shipboard Cycling System Diagnostics

by

**James P. Mosman**

B.S., Applied Mathematics and Physical Sciences, Colorado State University, 1997

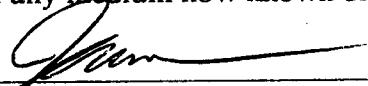
Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the  
Requirements for the Degrees of  
Naval Engineer  
and  
Master of Science in Ocean Systems Management  
at the  
Massachusetts Institute of Technology

June 2006

© 2006 James P. Mosman. All rights reserved.

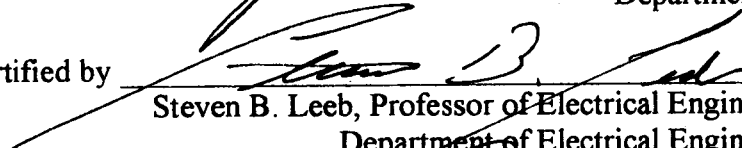
The author hereby grants to MIT and the US Government permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

Signature of Author



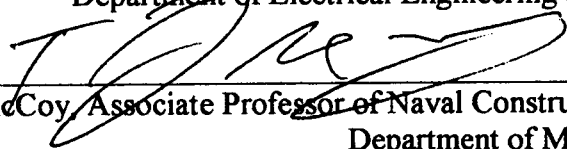
Department of Mechanical Engineering  
May 12, 2006

Certified by



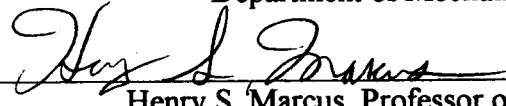
Steven B. Leeb, Professor of Electrical Engineering and Computer Science  
Department of Electrical Engineering and Computer Science  
Thesis Supervisor

Certified by



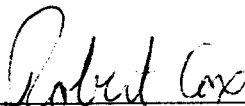
Timothy J. McCoy, Associate Professor of Naval Construction and Engineering  
Department of Mechanical Engineering  
Thesis Reader

Certified by



Henry S. Marcus, Professor of Marine Studies  
Department of Mechanical Engineering  
Thesis Reader

Certified by



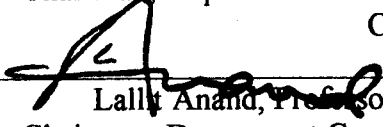
Robert W. Cox, Doctoral Candidate  
Department of Electrical Engineering  
Thesis Reader

Accepted by



Michael Triantafyllou, Professor of Mechanical Engineering  
Chairman, Department Committee on Graduate Students  
Center for Ocean Engineering

Accepted by



Lally Anand, Professor of Mechanical Engineering  
Chairman, Department Committee on Graduate Students  
Department of Mechanical Engineering

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE <b>01 JUN 2006</b>		2. REPORT TYPE <b>N/A</b>		3. DATES COVERED <b>-</b>	
4. TITLE AND SUBTITLE <b>Evaluation of Non-Intrusive Load Monitoring for Shipboard Cycling System Diagnostics</b>				5a. CONTRACT NUMBER <b>N62271-97-G-0026</b>	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S)				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) <b>Massachusetts Institute of Technology</b>				8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT <b>Approved for public release, distribution unlimited</b>					
13. SUPPLEMENTARY NOTES <b>The original document contains color images.</b>					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT <b>UU</b>	18. NUMBER OF PAGES <b>99</b>	19a. NAME OF RESPONSIBLE PERSON
a. REPORT <b>unclassified</b>	b. ABSTRACT <b>unclassified</b>	c. THIS PAGE <b>unclassified</b>			

# **Evaluation of Non-Intrusive Load Monitoring for Shipboard Cycling System Diagnostics**

by  
**James P. Mosman**

Submitted to the Department of Mechanical Engineering on May 12, 2006 in Partial  
Fulfillment of the Requirements for the Degrees of

Naval Engineer  
and  
Master of Science in Ocean Systems Management

## **ABSTRACT**

The Non-Intrusive Load Monitor (NILM) is a device that utilizes voltage and current measurements to determine the operating schedule of all of the major loads on an electrical service. Additionally, the NILM can use its electrical measurements to diagnose impending failures in the mechanical systems that are actuated by the electric loads. Ongoing NILM research conducted at Massachusetts Institute of Technology's Laboratory for Electromagnetic and Electronic Systems (LEES) is exploring the application of NILM technology in shipboard environments. For the current shipboard applications, diagnostic software development is in progress. To aid in that process, research was done to understand the dynamics of a shipboard cycling system.

This thesis presents an in-depth examination of the development of diagnostic indicators for a shipboard vacuum assisted waste disposal system. Measurements and experimentation were conducted onboard *USCGC SENECA (WMEC-906)*, a 270-foot Coast Guard Cutter. In order to better understand the system dynamics, a computer based model was developed to simulate the system. The intent of creating an in-depth model was to develop diagnostic methods that are applicable to any shipboard cycling systems.

First, a base model is designed followed by the exploration of a realistic model that includes variation commonly found in the system. Thirdly, a diagnostics section explores methods to detect increased pump operation and distinguish between high system usage and the presence of a leak. Lastly, a basic cost analysis is done on the sewage system to show the benefits of installing a NILM.

Thesis Advisor: Steven B. Leeb

Title: Associate Professor of Electrical Engineering and Computer Science

Thesis Reader: Timothy J. McCoy

Title: Associate Professor of Naval Construction and Engineering

Thesis Reader: Henry S. Marcus

Title: Professor of Marine Systems

## Acknowledgements

The author would like to acknowledge the following organizations and individuals for their assistance. Without them this thesis would not have been possible.

- The Office of Naval Research's Control Challenge, ONR/ESRDC Electric Ship Integration Initiative and the Grainger Foundation, all of whom provided funding
- LCDR Mike Obar, LTJG Jon Potterton, MK1 Labrier and EM1 Bassett for their support on USCGC Seneca
- Jim Paris and Chris Laughman for their computer and NILM technical assistance
- Rob Cox for his time and dedication to a Navy student's take on research
- CDR (ret) Timothy J. McCoy and Professor Henry S. Marcus for advisement as thesis readers
- Finally, Professor Steven Leeb who hunted me down a year ago and opened wide a door of opportunity that has been rewarding, challenging and, at times, entertaining

# Table of Contents

ABSTRACT.....	3
Acknowledgements.....	4
Table of Contents.....	5
List of Figures.....	7
List of Tables.....	10
1 Introduction.....	11
1.1 NILM Definition.....	11
1.2 Motivation for Research.....	12
1.3 Objective and Outline of Thesis.....	13
2 Basic Premises and Test Platform Description.....	15
2.1 NILM Basics.....	15
2.2 Cycling System Basics.....	18
2.3 USCGC SENECA Sewage System.....	20
3 Base Model System Characteristics and Simulation.....	25
3.1 Base Model System Assumptions.....	25
3.2 Basic Model Formulation.....	26
3.3 Building the Base Model Simulator.....	30
4 Real System Modeling.....	37
4.1 Real System Characteristics.....	37
4.2 Leak Rate Variation.....	38
4.3 Flush size variation.....	42
4.4 System Usage Rate Variation.....	45
4.5 Compilation of Variation in All Factors.....	47
4.6 Simulation Results.....	48
5 Diagnostic Indicator.....	51
5.1 Possible Diagnostic Methods.....	51
5.1.1 Mean Shift Test and Total Number of Pump Runs Test.....	51
5.1.2 Discontinuity Detection Test.....	52
5.1.3 Parameter Estimation and Trending.....	57
5.1.4 Parameter Trending.....	64
5.1.5 Load Time Analysis.....	71
5.1.6 Inport Data Considerations.....	72
5.2 Diagnostic Method and Status Reports.....	76
5.2.1 Inport Diagnostic Method.....	76
5.2.2 Underway Diagnostic Method.....	78
6 Cost Analysis for Monitoring of Seneca Sewage System.....	81
6.1 Motivation.....	81
6.2 Power Calculations.....	82
6.3 Cost-Benefit Analysis.....	83
6.4 Conclusions.....	86
7 Future Work and Conclusion.....	87
7.1 System Modeling.....	87
7.2 Diagnostic Indicator Testing.....	87
7.3 Cost Considerations.....	88
7.4 NILM Equipment.....	88

7.5 Conclusion .....	89
List of References .....	90
Appendix A: Simulink Model and Matlab Embedded Code .....	93
Appendix B: Matlab Code .....	97

## List of Figures

Figure 2-1: Line diagram of Non-Intrusive Load Monitor in a three phase electrical system. ....	15
Figure 2-2: Spectral envelopes recorded during the start of an incandescent lamp and an induction motor, respectively [10]. ....	16
Figure 2-3: Stator current (upper plot) and real power (lower plot) during the start of a vacuum pump motor. Overlaid atop the spectral envelope is a template that has been successfully matched to the observed transient pattern [10]. ....	18
Figure 2-4: Basic components of a typical cycling system. ....	19
Figure 2-5: USCGC Seneca (WMEC-906) and installed vacuum pumps. ....	20
Figure 2-6: USCGC SENECA sewage system basic schematic. ....	21
Figure 2-7: Seneca sewage system pressure trace (upper plot) and vacuum pump power (lower plot) chronologically aligned. The pressure decreases are caused by a system leak (the gradual decrease) and by toilet flushes (the step decreases). ....	22
Figure 2-8: Normal power traces for vacuum and discharge pumps. ....	23
Figure 2-9: Typical histogram of times between vacuum pump runs for seven day underway period (plot data from August 2005). ....	24
Figure 3-1: Base model pressure with no leak. Each line represents the possible system pressures. ....	27
Figure 3-2: Base model pressures with system leak. Each line represents possible system pressures. ....	28
Figure 3-3: Comparison of simulation results for various usage rates and leak rates. ....	31
Figure 3-4: One week simulation with no leak, $\lambda=30$ and Erlang of order 4 overlaid. ...	33
Figure 3-5: One week simulation with no leak, $\lambda = 30$ and Erlang of order 5 overlaid. .	34
Figure 3-6: One week of simulated data with 6 in-Hg/hour leak showing change in Erlang order. ....	35
Figure 3-7: Expected spike heights calculated from Erlang cumulative distribution function for spikes located at various times. Taken from ref [14] ( $\eta$ =Erlang order). ....	36
Figure 4-1: One week simulation baseline with 6 in-Hg/hour leak. ....	38
Figure 4-2: Simulated system pressure over time given that the vacuum pumps to not energize to raise pressure. ....	39
Figure 4-3: Simulated pressure trace in normal operating range of vacuum system. ....	40
Figure 4-4: Effect of varying leak rates on spike times. ....	41
Figure 4-5: One week simulation with leak rate variation. ....	42
Figure 4-6: Effect of varying vacuum loss due to a usage event. ....	44
Figure 4-7: One week simulation with usage event size variation. ....	44
Figure 4-8: One week simulation with three different lambda values. ....	46
Figure 4-9: One week simulation with compilation of parameter variations. ....	48
Figure 4-10: Comparison of Seneca underway data and simulated underway data for a seven day period with no leak. Seneca total number of runs = 1297, simulated total runs = 1288. ....	49
Figure 4-11: Comparison of Seneca data and simulated data for a five day period with 12 in-Hg/hr leak. Seneca total number of runs = 1102, simulated total runs = 1100. ..	49
Figure 4-12: Comparison of Seneca data and simulated data for a three day period with check valve failure. Seneca total number of runs = 1476, simulated total runs = 1463. ....	49

Figure 5-1: Histogram for three days of underway, no leak Seneca data overlaid with median filtered data (window size = 7). .....	53
Figure 5-2: Median filtering example on base model data. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data. ....	54
Figure 5-3: Median filtering example on five days of Seneca data with no leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data. ....	55
Figure 5-4: Median filtering example on five days of Seneca data with leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data. ....	55
Figure 5-5: Median filtering example on five days of Seneca data with check valve failure. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data. ....	56
Figure 5-6: Median filtering example on three days of inport Seneca data with large leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data. ..	56
Figure 5-7: Fitted Gamma distributions overlaid on Seneca data. ....	60
Figure 5-8: Chi-squared distribution with typical threshold shown. ....	62
Figure 5-9: Various chi-squared analysis methods on three weeks of simulated data. ...	63
Figure 5-10: Gamma distribution fitted $k$ values for varying leak rate (LR) and for varying usage rate (usage) using least squares, method of moments and maximum likelihood estimator methods. ....	65
Figure 5-11: Gamma distribution fitted $\lambda$ values for varying leak rate (LR) and for varying usage rate (usage) using least squares, method of moments and maximum likelihood estimator methods. ....	65
Figure 5-12: Trended $k$ and $\lambda$ values for simulated data containing two periods with a 12 in-Hg/hour leak present. ....	67
Figure 5-13: Trended $k$ and $\lambda$ values for simulated data containing two periods of elevated usage (increase of 10 flushes/hour). ....	68
Figure 5-14: The total number of pump runs for the simulated cases. ....	69
Figure 5-15: Total number of pump runs for Seneca data parameter trending periods. ..	70
Figure 5-16: Seneca $k$ and $\lambda$ values results for inserted leak cases. ....	70
Figure 5-17: Trended $k$ and $\lambda$ values for Seneca data with check valve failure. ....	71
Figure 5-18: Loaded run times for Seneca vacuum pumps with no leak condition (left) and 12 in-Hg/hour leak condition (right). ....	72
Figure 5-19: Inport Seneca data for no leak, 30 SCFH leak, 50 SCFH leak, and check valve failure. ....	73
Figure 5-20: Inport nighttime Seneca data for no leak, 30 SCFH leak, 50 SCFH leak, and check valve failure. ....	74
Figure 5-21: Photos of failed check valves: as opened (upper left), pitted valve face (upper right), rubber valve face with uneven wear marks (lower left), and pitted face of second valve (lower right). ....	75
Figure 5-22: Inport status report twenty-four hours after check valve failure. ....	77
Figure 5-23: Inport status report seventy-two hours after insertion of 17 in-Hg/hour leak. ....	78
Figure 5-24: Underway status report twenty-four hours after check valve failure. ....	79



Figure 5-25: Underway status report seventy-two hours after introduction of 12 in- Hg/hour leak.....	80
---	----

## List of Tables

Table 2-1: USCGC Seneca sewage system parameters and loads[15].	22
Table 3-1: Simulation inputs based on Seneca setpoints.	32
Table 3-2: Calculated "spike" times for a 6 in-Hg/hr leak and flush size of 1.1 in-Hg/flush.	35
Table 4-1: Expected spike times with variable leak	41
Table 4-2: Parameter variation allowed in the simulation model.	47
Table 5-1: Samples of means and total number of pumps runs.	52
Table 5-2: Parameter trending data arrangement scheme for simulated data for inserted leak and increased usage rates.	66
Table 5-3: Parameter trending data arrangement scheme for Seneca data for an inserted leak.	69
Table 5-4: Total number of runs and mean times between pump runs for Seneca inport periods.	73
Table 6-1: Average pump power level calculations.	82
Table 6-2: Excess energy costs associated with a 30-day undetected check valve failure condition.	83
Table 6-3: Inputs used in cost-benefit model to determine lifetime costs of shipboard sewage system.	85
Table 6-4: Discounted lifetime costs for shipboard sewage system with and without NILM installed.	86

# 1 Introduction

## 1.1 NILM Definition

The Non-Intrusive Load Monitor (NILM) is a device that utilizes electrical voltage and current to determine the operating schedule of major loads. The non-intrusive aspect of the device is its minimal impact on an existing system. Simple wire connections are used to monitor the voltage and a current transducer is used to measure the aggregate current. These raw measurements are analyzed by the installed software to calculate the real and reactive power which in turn can be used to perform diagnostics on the electrical system.

Non-intrusive load monitoring research has been conducted at Massachusetts Institute of Technology's Laboratory for Electromagnetic and Electronic Systems (LEES) over the past two decades. The NILM has been previously used in residential, commercial and automotive environments [1][2][3]. The research presented in this thesis is for the application of NILM technology in shipboard environments. Previous research has shown the NILM to have potential in this environment and warrants further research and development.

For the current shipboard applications of NILM, the transient event detection and diagnostics software has yet to be fully written. To aid in the development of the NILM software, research is necessary to understand dynamics of the shipboard system. The research presented in this thesis is an in-depth examination of the development of diagnostic indicators and leak detection methods for a shipboard cycling system. In order to better understand the system dynamics, a computer based model is developed to simulate the system and better test the diagnostic methods. The goal of exploring the model development step-by-step is to make this method applicable to any shipboard cycling system. A basic cost analysis of the advantage of using a NILM is also done for a specific system onboard the test ship.

## 1.2 Motivation for Research

Electrical components have been onboard ships since the 1800s. Since the first application, the population of electrical components onboard has only increased. The development of the computer and modern microelectronics has greatly increased the demand for electrical generation and has also increased the complexity of the systems. Today, electrical components are integral in every system onboard a ship. Electrical systems have become the single-most important system on any ship. In the near future, electricity will likely become the primary source for propulsion power as well as provide the propulsive force in advanced weapons systems.

Electrical components are not only stand-alone, such as a gun control system, but are also components of mechanical systems, such as a pump in a seawater cooling system. Since all systems require some amount of attention, the users of ship systems must be able to determine the status or condition of a system at any time. Traditionally, the monitoring has been done with watchstanders taking logs and with dedicated sensors whose outputs are input into a larger monitoring circuit. These sensors are often intrusive in that they must break system integrity to monitor such characteristics as pressure or temperature. Large systems can have many sensors which require complex monitoring circuits. A typical engine room onboard a modern Navy warship can have hundreds to thousands of sensors. Nearly all the sensors monitor only one system parameter and often have redundant sensors in the same system to improve monitoring reliability. As more automated engine rooms are designed for new warships, the number of sensors has the potential to increase nearly two orders of magnitude [8]. With the increase in sensors comes an increased amount of wiring, complexity, weight, and cost. Shipboard NILM installations have the potential to avert those increases and reduce shipbuilding costs.

Although current and voltage are currently monitored on some systems, it's usually done to check for overcurrent and over/undervoltage conditions. The NILM uses only these two inputs to perform its analyses and is connected at a single point. A majority of mechanical systems have electrical components whose operation not only depends on the component itself, but also the mechanical system to which it is attached. The NILM concept applied to shipboard systems uses only electrical power to determine the health of an electro-mechanical system. Single-point monitoring of the electrical

power has the potential of informing the user of the overall health of a system and reducing the need for extra sensors and monitors.

### 1.3 Objective and Outline of Thesis

The research presented in this thesis is a continuation of research conducted by LCDR Jack S. Ramsey, Jr., USN [6] and by LT Thomas W. DeNucci, USCG [7]. In LCDR Ramsey's thesis, the feasibility of using NILM was tested on multiple shipboard systems onboard three different ships. His results were positive and he concluded that the NILM could be used successfully in the shipboard engineering environment. LT DeNucci's thesis explored diagnostic indicators for shipboard cycling systems, diagnostic indicators of a pump-motor coupling failure, analyses of fluid system blockages, and analyses of NILM applications on a reverse osmosis system. LT DeNucci's results were also very promising and he concluded that NILM could be used to diagnose pathological equipment failures.

The purpose of this thesis is to further explore and develop the diagnostic indicators for a shipboard cycling system. An in-depth analysis of the cycling system is presented and a realistic model is created to accurately simulate the cycling system. Although the research presented is for one specific cycling system onboard one ship, the methods used are intended to be applicable to any cycling system on any ship. Chapter Two discusses some NILM and cycling system basics and describes the test platform. Chapter Three discusses the development of a simulation model for an ideal cycling system. Chapter Four enhances that model by adding realistic dynamics into the simulation. Chapter Five discusses the diagnostic indicators for the cycling system. Chapter Six presents a basic cost analysis of a situation where no monitoring was done on the cycling system, and Chapter Seven presents recommendations, future work and conclusions.

Page Intentionally Left Blank

## 2 Basic Premises and Test Platform Description

### 2.1 NILM Basics

A line diagram of a NILM system hooked to a three phase electrical system is shown in Figure 2-1. The NILM concept is based on the observation that the transient behavior of an electrical load is influenced by the task that the load performs [4]. As a result, different loads possess unique and repeatedly observable transient profiles which can serve as “fingerprints” associated with each load. One example of this difference is a comparison of the turn-on transients associated with an incandescent lamp and an induction motor as shown in Figure 2-2. The physical task of heating a cold lamp filament is unique from the acceleration of a rotor [4]. The NILM was developed to detect the operation of individual loads using transient patterns observed in the short-time estimates of the spectral content of the aggregate current drawn by a collection of loads [4][5].

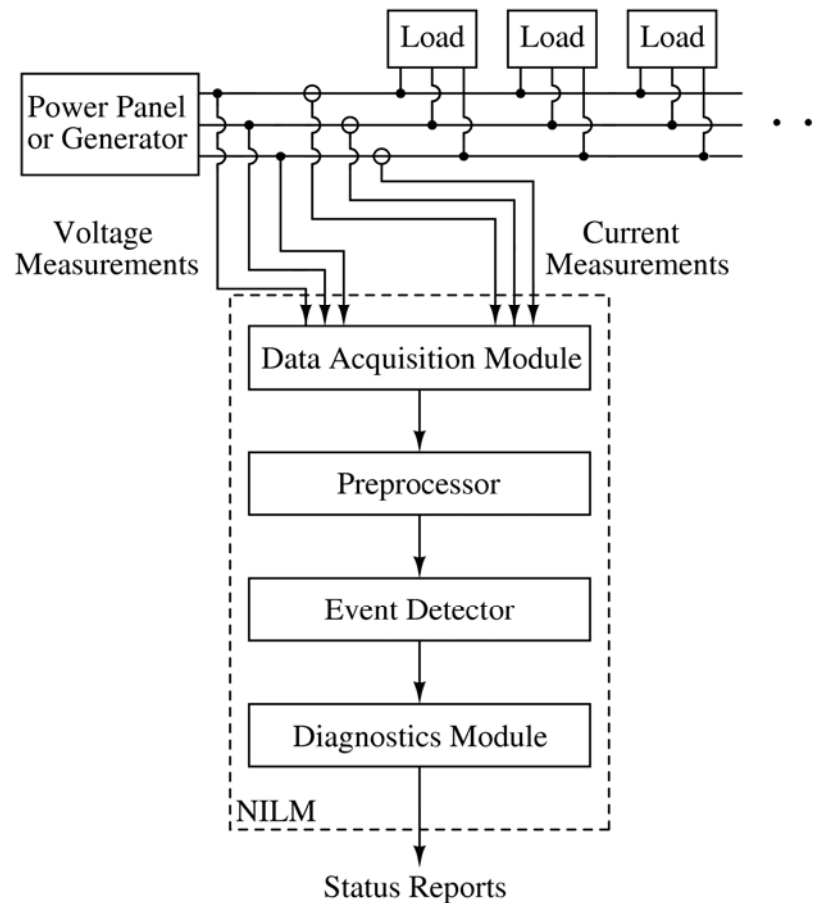
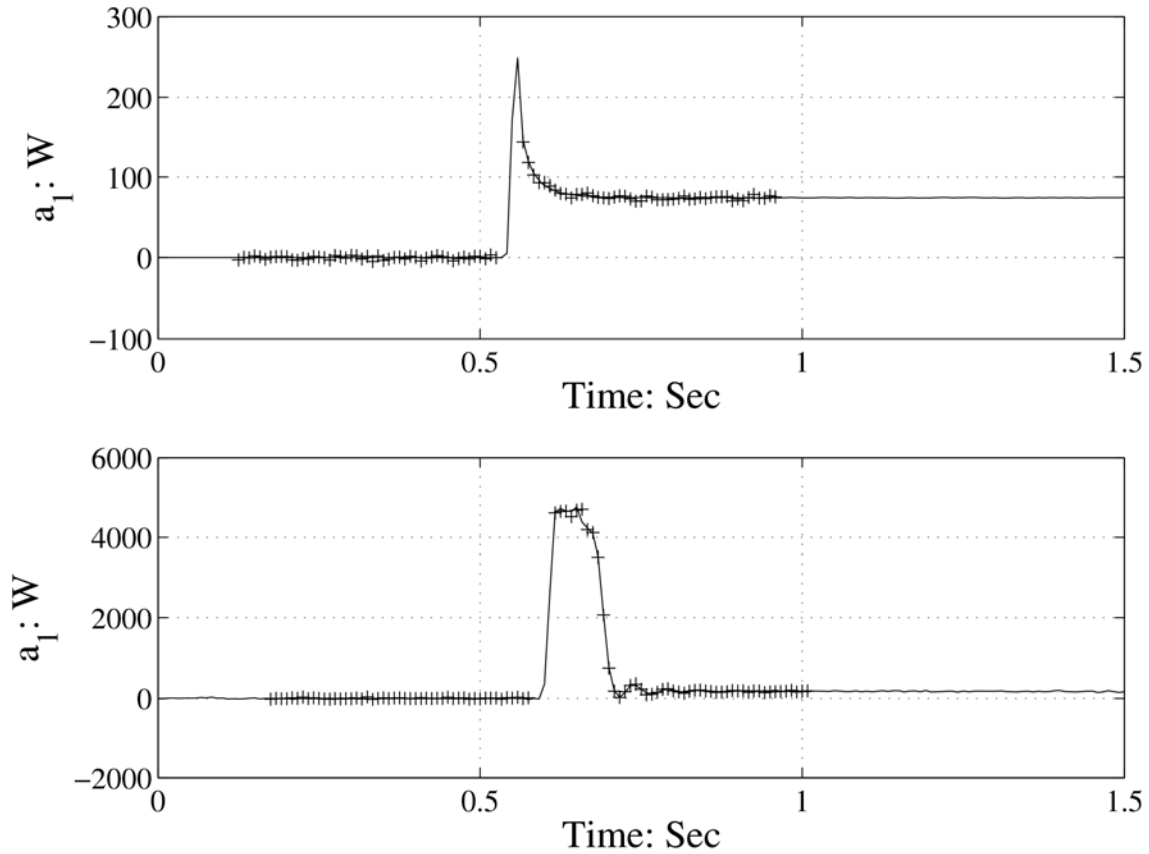


Figure 2-1: Line diagram of Non-Intrusive Load Monitor in a three phase electrical system.



**Figure 2-2: Spectral envelopes recorded during the start of an incandescent lamp and an induction motor, respectively [10].**

As shown in Figure 2-1, the NILM system uses single point voltage and current measurements to estimate real and reactive power. The NILM does not interfere with the load(s) downstream of the measurement point. A NILM setup consists of a Pentium class PC, a data acquisition card, a keyboard and monitor for user interface, a NEMA-style box to house the sensing boards and a power supply board, and the associated wiring to connect the NILM to the sensors and to the power supply.

The voltage sensing connection, external to the NEMA box, is a wired connection from the ship's power panel to the voltage sensing board inside the NILM setup. Current sensing is done using a commercial off-the-shelf (COTS) current transducer placed around each of the phases leading to the load(s) fed by that power supply. Although Figure 2-1 shows connection to all three phases of voltage and current, only two phase voltages and one phase current in an ungrounded three phase system are required for



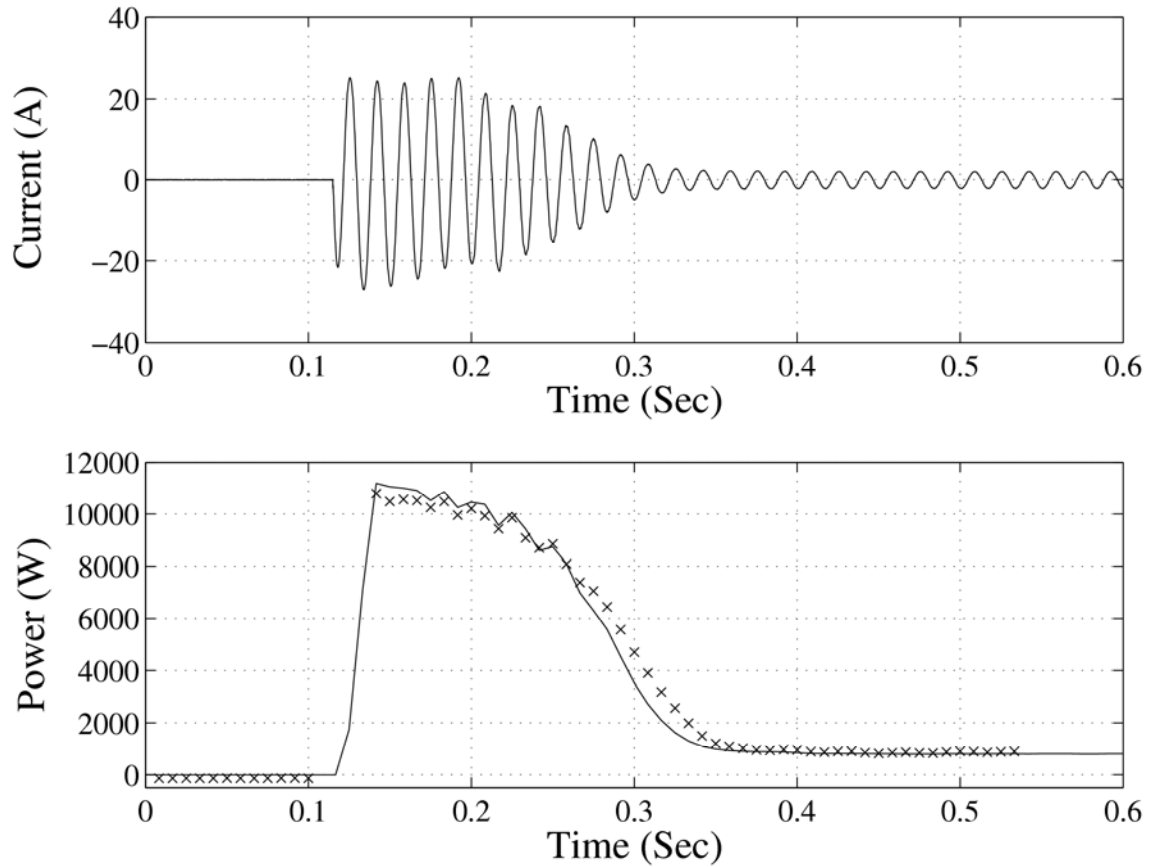
NILM operation. A more detailed description of the components and how they are connected is available in reference [6].

In order to accurately monitor short electrical transients, a relatively high (8 kHz) voltage and current sampling rate is used to capture data and the resulting power envelope data rate is 120 Hz[4][5]. Spectral envelope coefficients, defined in equations (1.1) and (1.2), contain time local information about the frequency content of  $x(t)$ . The spectral envelope equations are Fourier-series analysis equations evaluated over a moving window of length  $T$  where  $m$  is an integer and  $\omega$  is the base frequency. In a steady-state AC power system like that onboard a ship, the spectral envelope coefficients have a useful physical interpretation as real power, reactive power, and harmonic contents when  $x(t)$  is the measured current and the sine and cosine terms are synchronized with the voltage [9].

$$a_m(t) = \frac{2}{T} \int_{t-T}^t x(\tau) \sin(m\omega\tau) d\tau \quad (1.1)$$

$$b_m(t) = \frac{2}{T} \int_{t-T}^t x(\tau) \cos(m\omega\tau) d\tau \quad (1.2)$$

For the applications used in this thesis, only the real power was utilized. Figure 2-3 shows the actual stator current, which is input to the NILM, and the real power, which is a NILM output, for a start of a vacuum pump motor. Overlaid on the lower plot is a “fingerprint” template that has been successfully matched to the pump start transient and thus can be used to identify the start in a transient event detector.



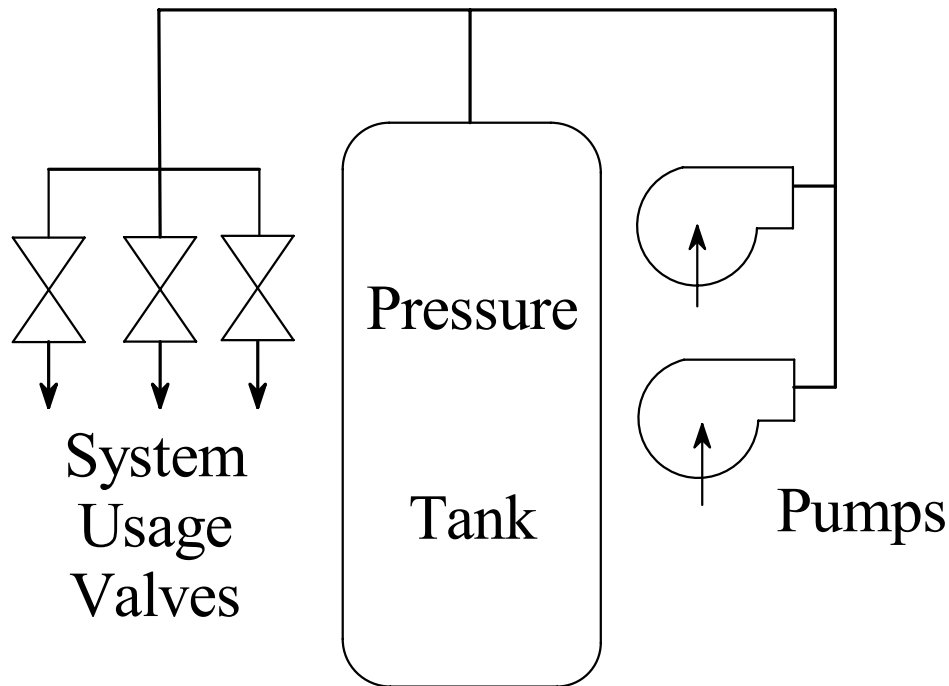
**Figure 2-3: Stator current (upper plot) and real power (lower plot) during the start of a vacuum pump motor. Overlaid atop the spectral envelope is a template that has been successfully matched to the observed transient pattern [10].**

At this current time and stage of NILM development, the transient event detector and diagnostics module are not fully developed, so the files are sent directly into data storage. The LINUX-based software included in the NILM can easily be updated to include transient event detection and diagnostic software. Research done in this thesis aids in the further development of NILM software and is intended for immediate implementation.

## 2.2 Cycling System Basics

Cycling systems are usually comprised of a capacitive element, a method of “recharging” the system and paths of energy release. As shown in Figure 2-4, the typical cycling system seen onboard a ship contains a tank, pumps to recharge the tank and piping with valves leading to other systems which draw fluid from the tank. Examples of

such systems include pressurized air systems and potable water systems where the pumps provide the air or water to a tank and the rest of the system draws from the tank through system usage valves.



**Figure 2-4: Basic components of a typical cycling system.**

Another cycling system which works on the same principle but is slightly different is a vacuum assisted drainage collection system. The pumps draw a vacuum on the tank and the rest of the system feeds into the tank. Essentially, the arrows and flow paths are reverse of what is shown in Figure 2-4. In this case, the vacuum pressure is stored by the tank and used by the rest of the system.

The next chapter will investigate the operation and characteristics of a base model of one such system. The model system is based on an actual system found onboard a U.S. Coast Guard cutter. Understanding the underlying dynamics of the system is important to understand how to model the system and develop diagnostic indicators. The fourth chapter will investigate the real system dynamics and how variance in parameters affects the results found from the base model situation.

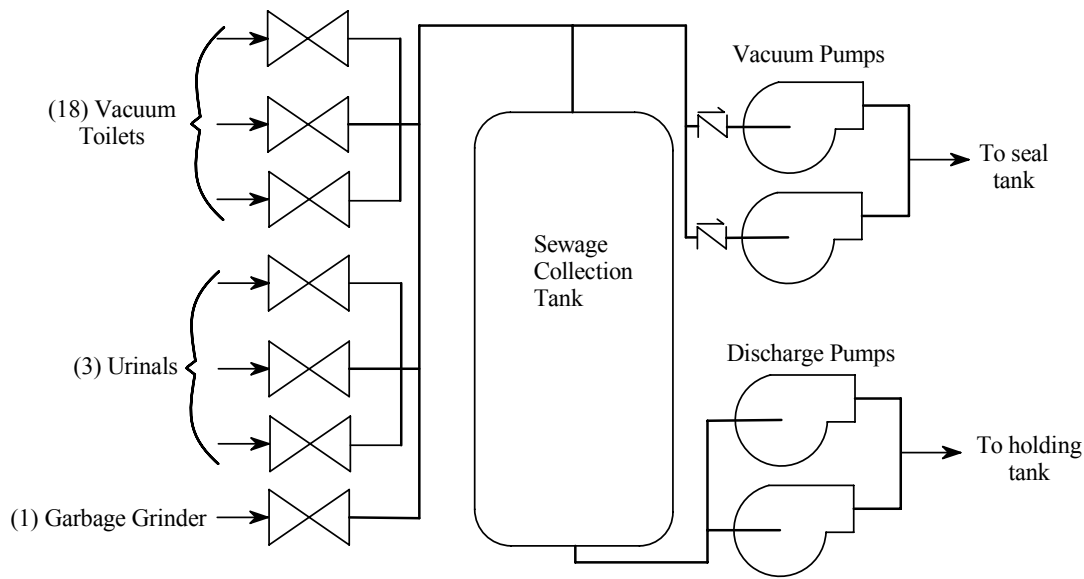
## 2.3 USCGC SENECA Sewage System

The ability to conduct tests and collect data on an active duty ship platform is essential to the success of the shipboard NILM project. The U.S. Coast Guard Cutter Seneca (WMEC-906) is the sixth of thirteen Famous Class medium endurance cutters. The ship's primary missions are to assert effective Search and Rescue (SAR) and Maritime Law Enforcement (MLE) in domestic or foreign waters. The ship has a length of 270 feet and displaces 1850 tons [11]. Figure 2-5 shows a recent picture of USCGC Seneca [11].



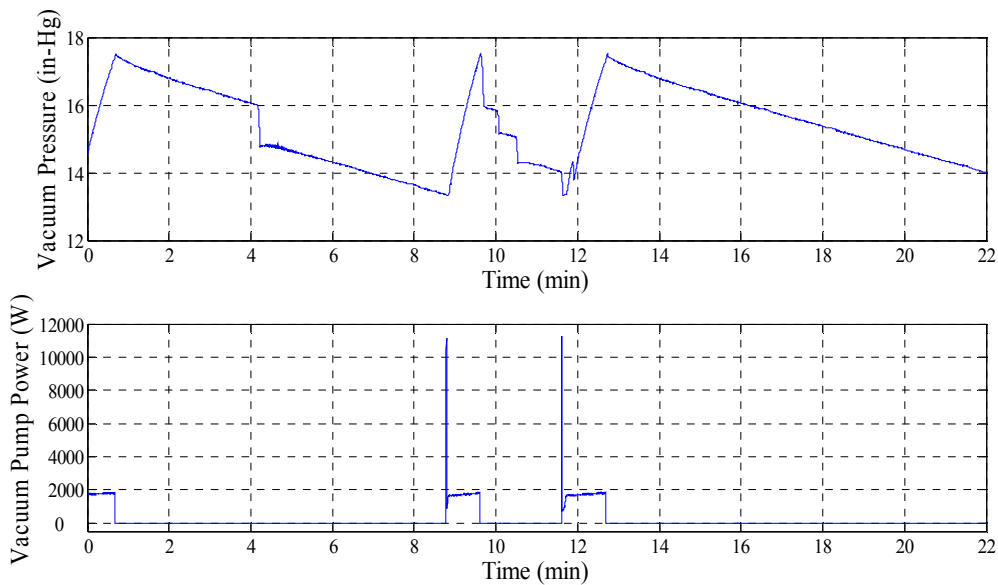
**Figure 2-5: USCGC Seneca (WMEC-906) and installed vacuum pumps.**

The system being studied and modeled is a vacuum assisted sewage collection system. The tank and pumps are located in an auxiliary machinery space onboard the ship. The system receives the drains from eighteen vacuum toilets, two urinal lift valves, one urinal non-lift valve and one galley garbage grinder. A 360 gallon collection tank stands upright with two 1.5 HP vacuum pumps connected to the top of the tank via piping and two check valves that function to retain the system vacuum pressure when the pumps are deenergized. The toilets, urinals and garbage disposer are zoned throughout the ship and lead into the top of the tank through isolation valves. A separate tank discharge system with two 2.0 HP pumps automatically drains the collection tank based on tank level [15]. Figure 2-5 contains a photo of the vacuum pumps and the holding tank and Figure 2-6 shows a basic system schematic.



**Figure 2-6: USCGC SENECA sewage system basic schematic.**

The vacuum pumps operate to maintain vacuum in the system. When the system pressure drops to 14 in-Hg, one vacuum pump energizes. Consecutive starts alternate between pumps to equalize the wear. If the pressure drops to 12 in-Hg, the second vacuum pump starts to assist the already running pump. The pump(s) de-energize when the tank pressure reaches 18 in-Hg [15]. Figure 2-7 shows the relationship between the vacuum pump power and the system pressure. The pressure data and pump run data was taken simultaneously during a leak period and aligned chronologically for comparison. Note that actual setpoints in the system are approximately 0.5 in-Hg lower than described in the tech manual. The smaller “down-steps” in middle of the traces correspond to usage events such as toilet flushes.



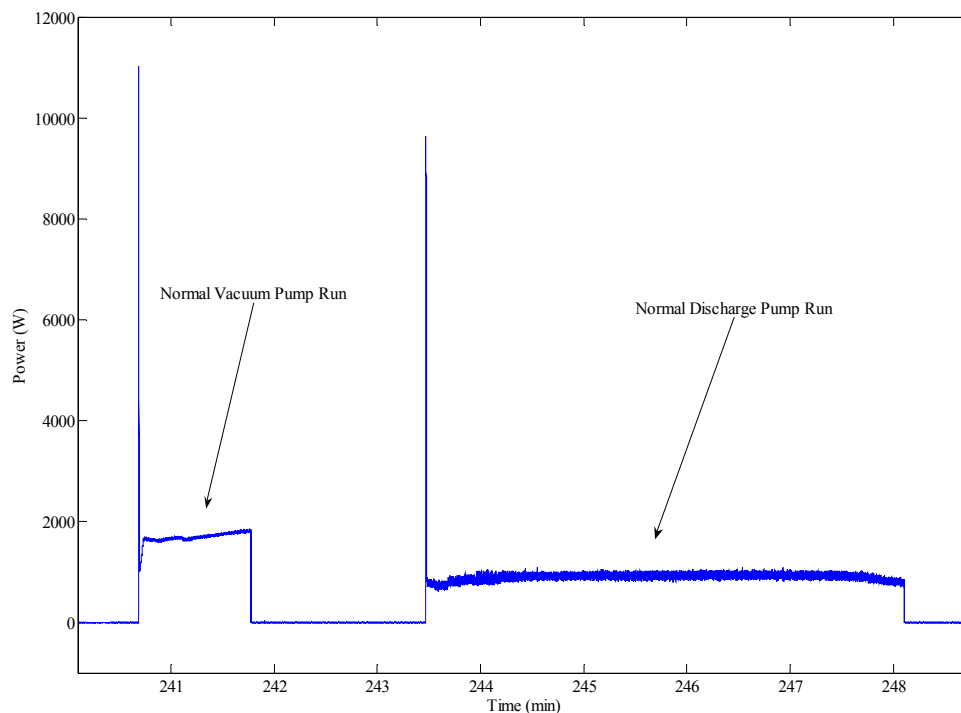
**Figure 2-7: Seneca sewage system pressure trace (upper plot) and vacuum pump power (lower plot) chronologically aligned. The pressure decreases are caused by a system leak (the gradual decrease) and by toilet flushes (the step decreases).**

The discharge pumps energize, alternating on consecutive starts, when the water and waste level in the tank reaches 33% of its full capacity (120 gallons) and de-energize when the level is 5% (18 gallons). Water and waste is pumped from the vacuum system to an atmospherically pressured holding tank for later discharge overboard or to a collection system on the pier. Table 2-1 lists the system setpoints, pump capacities, and system loads [15].

**Table 2-1: USCGC Seneca sewage system parameters and loads[15].**

<u>Parameter</u>	<u>Value</u>
High Vacuum ( $P_0$ )	18 in-Hg
Low Vacuum ( $P_{low}$ )--1 pump starts	14 in-Hg
Lower Vacuum ( $P_{lower}$ )--2 pumps start	12 in-Hg
Vacuum Pump Capacity (each)	23 cfm @16 in-Hg
Discharge Pump Capacity (each)	30 gpm
Holding tank capacity	360 gallons
System capacity (approx.)	600 gallons
<u>System Loads</u>	
(18) Vacuum Toilet Assemblies	≈0.375 gal per flush
(3) Vacuum Urinal Assemblies	≈0.25 gal per flush
(1) Garbage Grinder Kit	≈0.83 gal per use

The NILM was installed in the control panel for the vacuum and discharge pumps by Ramsey in 2003 [6]. Two phases of the 440 volt electrical power in the pump controller are measured and the current is measured on the third phase. Both the vacuum pumps and discharge pumps use the same power supply so their input voltages are the same. The current transducer was installed to measure the current passing to the four pumps collectively. That is, if both vacuum pumps were energized and one of the discharge pumps energized, the current sensed would be the sum of the currents to the three individual loads. A typical power plot showing both a vacuum pump and the discharge pump is show in Figure 2-8.

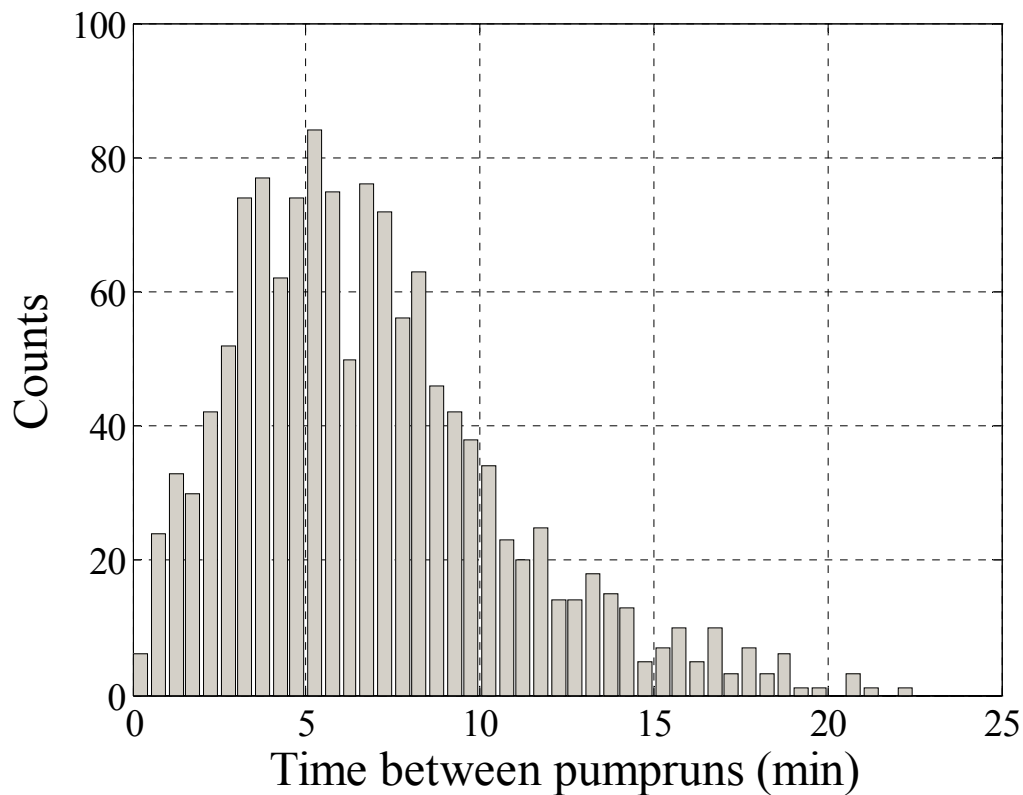


**Figure 2-8: Normal power traces for vacuum and discharge pumps.**

Of primary interest to the author were the effects of increased system usage and system leaks on the frequency of vacuum pump runs and how each of the system characteristics affected the dynamics of the entire system. The goal of the research was to be able to determine the normal operating conditions of the systems and to be able to diagnose the presence of a leak in the system. Since the frequency of vacuum pump runs is the directly related to the usage of the system and the presence of any vacuum leaks,

the focus of the research was time between vacuum pump runs. The discharge pump runs were largely ignored in the data analysis, but mention of their importance will be discussed later with respect to creating a diagnostic indicator.

Data collected by the NILM was analyzed using MATLAB scripts to detect the times between the securing of one vacuum pump and the start of the next vacuum pump. The collected times between pump runs were then binned in a histogram with equal bin sizes in order to give a display of the system usage. Figure 2-9 below shows a typical histogram of the times between pump runs for an underway period of five days.



**Figure 2-9: Typical histogram of times between vacuum pump runs for seven day underway period (plot data from August 2005).**



### 3 Base Model System Characteristics and Simulation

#### 3.1 Base Model System Assumptions

Since underway experimentation time onboard the Coast Guard cutter was limited, a computer based model was needed in order to better understand the system characteristics and to produce data for development of a diagnostic indicator. In order to develop the model, each factor that influenced the system needed to be explored and understood. The remaining portions of this chapter will discuss the formation of a base model with no parameter variation and predictable results.

A system usage event is caused by the crew flushing a toilet, flushing a urinal, or using the garbage disposer. Discussions with the crew revealed that the garbage disposer is not operated very often, so the flushing events are the primary influences on the system. An “event” was defined as one flush of a toilet or urinal. The crew flushing behavior was investigated by DeNucci and most closely resembles a naturally occurring Poisson process [7]. For a Poisson process, the time between the  $k^{\text{th}}$  event and the  $(k-1)^{\text{th}}$  event can denoted by a random variable  $T_k$ , is alternately referred to as the  $k^{\text{th}}$  inter-arrival time and is distributed according to the following probability density function (PDF) [14].

$$f_{T_k} = \lambda e^{-\lambda t} \quad (3.1)$$

Given this hypothesis, the crew usage rate,  $\lambda$ , has a direct effect on the measured times between pump runs. More flushes results in more vacuum loss and thus an increased frequency of pumps runs to recharge the vacuum tank.

Another vacuum pressure reduction factor is the size of a system usage event. The amount of vacuum lost during one flush of a toilet or urinal also directly effects the times between pump runs. Larger flush drops result in more pump runs in a given period of time.

A third factor that affects the times between pump runs is the presence of a vacuum leak in the system. For obvious reasons, a larger leak rate results in increased pump run frequency.

To simplify the system and study the effects of each one of the above factors, assumptions had to be made. For the bose model system, the following assumptions were made:

- Every flush instantaneously removes the same amount of vacuum from the sewage system
- The leak rate is constant regardless of the system pressure
- Flushes occur according to a Poisson process and at a constant rate,  $\lambda$

Reasons for these simplifications will be explained in the following sections. The next chapter will explore deviation from these assumptions and the effects on the data received.

The controlling parameter in the simulation is pressure. Similarly to the real system, the pressure determines when the pumps are running and when they shut off. The vacuum pressure in the system is measured in in-Hg where the “high” vacuum pressure is actually the lowest absolute pressure. To avoid confusion, the simulation and the following discussions are done entirely in in-Hg hence the term “pressure” is synonymous with “vacuum pressure.”

## 3.2 Basic Model Formulation

There are two loss mechanisms that will reduce the system pressure. A flush, or a system usage event, will reduce the pressure by a discrete amount and a leak in the system will reduce the pressure as a function of time. Given these two loss mechanisms a basic, linear approximation of pressure can be written as

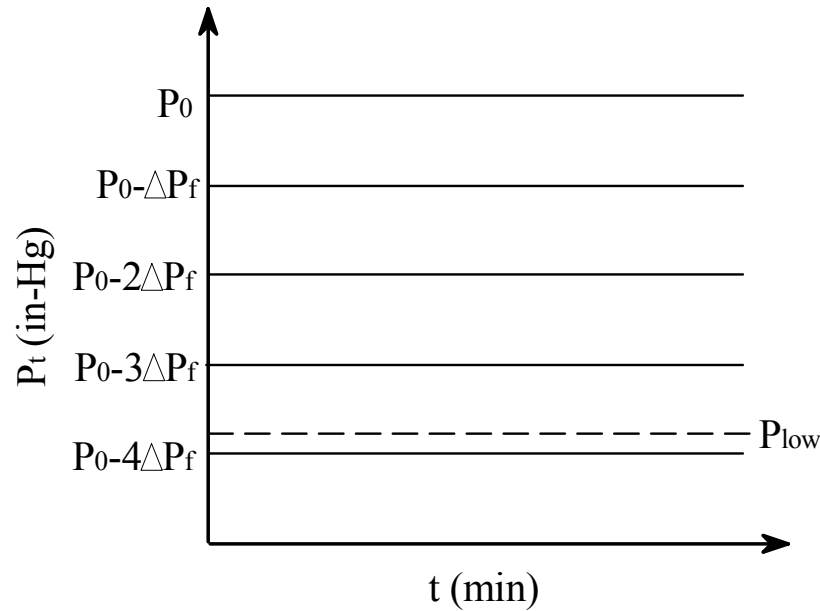
$$P_t = P_0 - N(\Delta P_f) - t \alpha_{leak} . \quad (3.2)$$

The variable  $P_t$  is the system pressure at time  $t$ ,  $P_0$  is the high pressure set point when the pumps turn off,  $N$  is the number of flushes which have occurred up to time  $t$ ,  $\Delta P_f$  is the amount of vacuum removed by a single flush, and  $\alpha_{leak}$  is the rate at which the leak reduces pressure.

The effects of each loss mechanism can be investigated by setting the other to zero. First, to examine the effects of flushing, the leak rate  $\alpha_{\text{leak}}$  is set to zero, so equation (3.2) becomes

$$P_t = P_0 - N(\Delta P_f) . \quad (3.3)$$

The first assumption introduced in section 3.1 is required in order to make this the base model. Variation in the flush size would eliminate the discreteness of the pressure values and complicates the evaluation. Later evaluation in the next chapter shows how flush size variation affects the results. Figure 3-1 shows the possible pressures at any time  $t$ . The  $t=0$  point corresponds to the time at which the vacuum pumps de-energized upon reaching the high pressure setpoint. The range between  $P_0$  and  $P_{\text{low}}$  depends on the system setpoints and  $\Delta P_f$  depends on the characteristics of the toilet or urinal being flushed.



**Figure 3-1: Base model pressure with no leak. Each line represents the possible system pressures.**

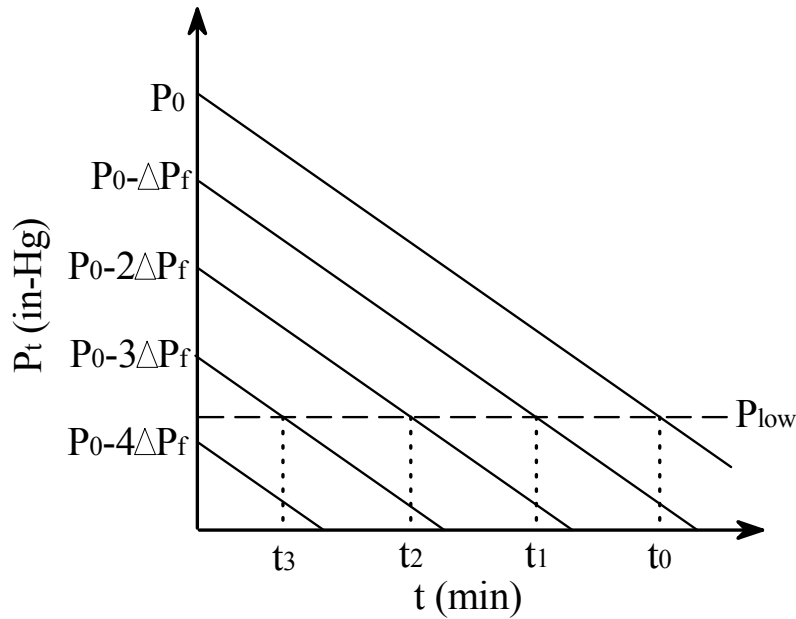
In a base model system with no leaks, the pressure reached after a pump operation would discretely decrease in even steps until the pressure in the system was at or below the low pressure setpoint and the pumps would reenergize to raise system pressure again.

It can be seen in the Figure 3-1 and derived from equation (3.3) that the number of flushes,  $N_{\max}$ , required to reach the low pressure setpoint is

$$N_{\max} = \left\lceil \frac{P_0 - P_{\text{low}}}{\Delta P_f} \right\rceil, \quad (3.4)$$

where  $\lceil \cdot \rceil$  is the ceiling function. Since it is impossible to have fractions of flushes, the ceiling function is used. It is important to note that with no leak in the system,  $N_{\max}-1$  flushes can occur without the pumps energizing.

To further develop the base model, the effect of a leak can be included and the second assumption from section 3.1 is enforced. From (3.2), it can be seen that the leak linearly decreases the pressures as time progresses with the effects of the number of flushes and size of flushes remaining the same. The size of the leak is assumed to remain constant regardless of system pressure so that the slope of the line remains linear. The effect of system pressure dependent leak rates is discussed in Chapter 4. The result of the base model with a constant leak rate and unique flush sizes is shown in Figure 3-2.



**Figure 3-2: Base model pressures with system leak. Each line represents possible system pressures.**

Given that a pump will energize when the pressure has dropped to the low pressure setpoint, the expected times between pump runs can be determined. The times at which  $P_{low}$  is reached can be derived from equation (3.2) by setting  $P_t$  equal to  $P_{low}$  and mathematically written as:

$$t_N = \frac{P_0 - N(\Delta P_f) - P_{low}}{\alpha_{leak}}, \quad (3.5)$$

where  $t_N$  is the expected times for a pump to energize given  $N$  flushes. This is demonstrated graphically in Figure 3-2 for varying numbers of flushes. Note that  $t_0$  is the longest time and  $t_N$  is the shortest. If no flushes had occurred ( $N=0$ ), the only effect on the system would be the pressure drop due to the leak and the expected time between pump runs would be  $t_0$ . Likewise, if  $N_{max}-1$  flushes had occurred, then the expected time between runs would be  $t_{N_{max}-1}$ .

For large leaks, the effect is a steeper slope of pressure lines and vice versa for a small leak. This intuitively makes sense with the no flush scenario. For a large leak, the time until a pump starts will be less than that for a small leak because the leak takes less time to deplete the vacuum in the system.

The pressure at time  $t$  and ultimately the time between pump runs is largely determined by how often the system is used, or, in other words, how often a drop in pressure occurs due to a flush. Previous research has examined how the crew behavior can modeled as a Poisson process [7]. Poisson processes require time homogeneity, meaning that the probability of  $k$  arrivals is the same for all time intervals of the same length, and they require independence, meaning that the number of arrivals in one time period is independent of the history of arrivals outside that time interval [12]. Both of these requirements are assumed for the base model and are included in the third base model assumption in section 3.1.

An observation was made by DeNucci that the inter-arrival times between flushes were exponentially distributed and thus led to an Erlang distribution of times between pump runs. The Erlang PDF is mathematically represented as

$$f_{erl}(k, \lambda, t) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!}, \quad (3.6)$$

where  $\lambda$  is the system usage rate (in flushes/hour),  $k$  corresponds to the  $k^{\text{th}}$  arrival, and  $t$  is the time elapsed (in hours) since the de-energization of the vacuum pumps.

The time at which the vacuum pumps de-energize after “recharging” the system is essentially a renewal event. This means that the pressure reduction process is restarted each time and the past system history does not affect the current pressures. A Poisson process depends on a renewal event to restart each process. The real system is slightly different in the fact that a flush might have occurred a few seconds prior to the pumps deenergizing, and the next flush might occur a few seconds after the pumps de-energize. A true renewal event means that the flush that occurred prior to the pumps deenergizing would not matter and time would “reset” to zero when the pumps shut off. This problem of the real system not having a true renewal event will be discussed in the next chapter, however for the base model examined here, it is assumed that each pump shut off is true renewal event and thus the Poisson process starts over at  $t=0$  each time.

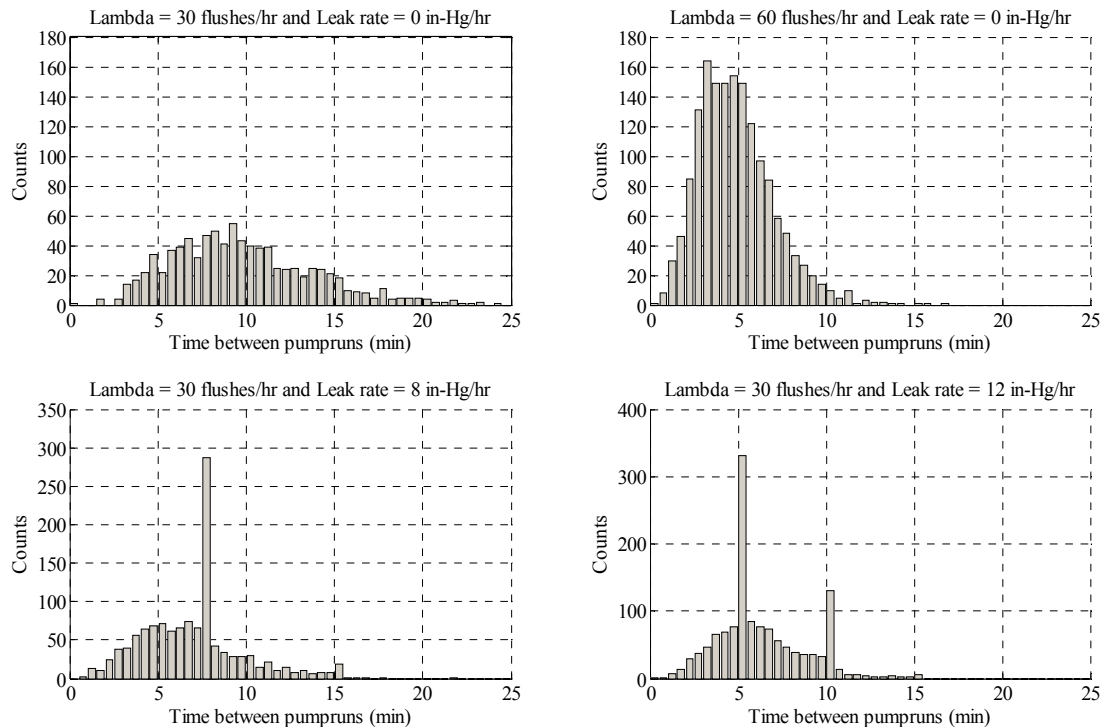
An Erlang probability density function arises when examining the inter-arrival times (times between flushes). Written out, the Erlang PDF translates to the probability that the  $k^{\text{th}}$  arrival will fall between times  $t$  and  $t+\Delta t$  and is equal to the probability that  $k-1$  arrivals have occurred in  $[0,t)$  multiplied by the probability that one more arrival will occur in time  $\Delta t$ . Observing Figure 3-2, it can be seen that for a value of  $k$ , the above probability relation is correct and appropriately applies to the cycling system. If time  $t_k$  is reached without the pump running, then the chance that a pump will run before  $t_{k-1}$  depends on the probability that  $k-1$  flushes have already occurred and the probability that a flush will occur before  $t_{k-1}$ .

### 3.3 Building the Base Model Simulator

Now that the effects of the base system characteristics are known, a model can be built to simulate that system. A model was built using MATLAB and Simulink. The foundation of the model is equation (3.2) where a discrete time simulation was developed using the linear relationships. A “prep” file, include in Appendix B, was created to develop a list of times at which flushes would occur using the MATLAB coding techniques introduced by DeNucci [7]. The times for the “prep” file are dependent on  $\lambda$  and the length of time simulation.

The model uses a summing function to analyze the pressure at each time step. An adequate time step used was one second. Each simulation second, the model sums the negative effect of a leak ( $=\text{leak\_rate} * \text{time\_step}$ ), the negative effect of a flush (if one had occurred in the last second) and the positive effect of a pump or pair of pumps running ( $=\text{pump\_rate} * \text{time\_step}$ ) on the system pressure. A logic routine that observes current pressure and determines how many pumps should be running is used to determine the “pump\_rate” used.

The output of the model is a vector with time in one row and a series of zeros and ones in the other row. A “0” indicates that no pumps are running and a “1” indicates that one or two pumps are running. The vector is sent to a “post” routine, included in Appendix B, that measures the time between pump runs and displays the results in a histogram. Outputs of this base model simulation are shown in Figure 3-3 below for varying levels of  $\lambda$  and for varying leak rates.



**Figure 3-3: Comparison of simulation results for various usage rates and leak rates.**

As can be seen in the upper plots of Figure 3-3, the  $\lambda$  value greatly affects the shape of the curve. A larger  $\lambda$  value means that the crew is using the system more often, so the mean time between runs should decrease and the total number of runs in a given

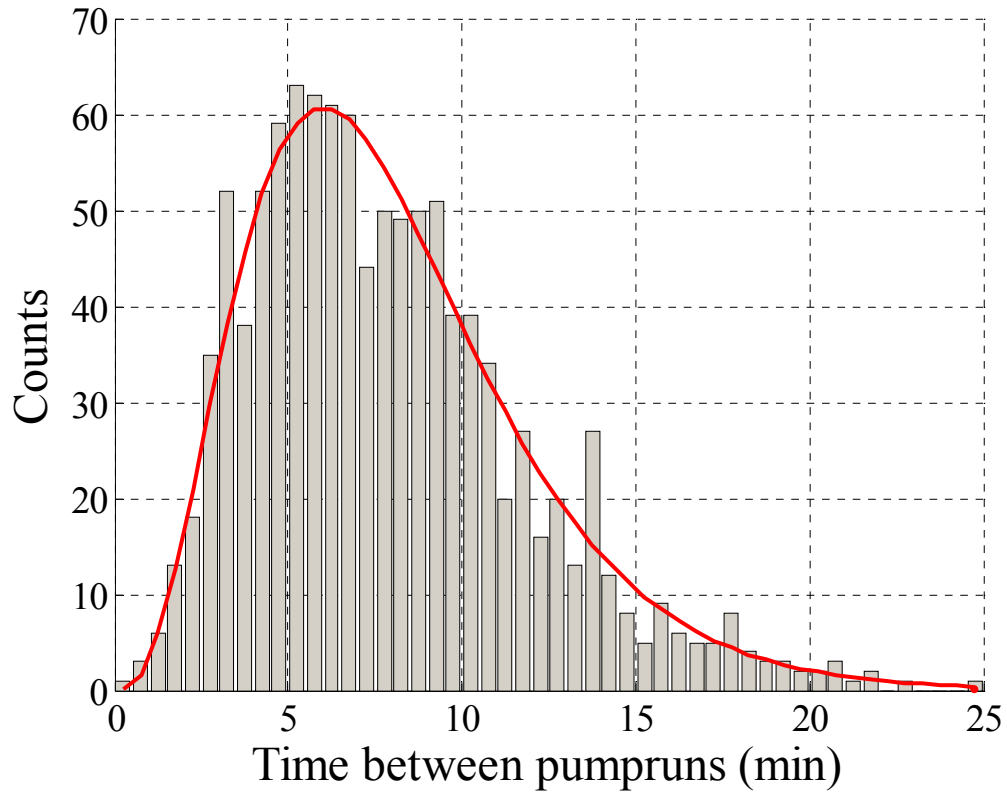
time period should increase. Another factor that affects the shape but is not as obvious is the size of the flush. The amount of vacuum removed by one flush,  $\Delta P_f$ , as seen in Figure 3-1 determines how many flushes are required to reach  $P_{low}$  and energize the vacuum pump. This flush size ultimately determines the “k” value in equation (3.6). To demonstrate this fact, consider the following setpoints input into the simulation.

**Table 3-1: Simulation inputs based on Seneca setpoints.**

<u>Parameter</u>	<u>Value</u>
Elapsed time	1 week
Leak rate	0 in-Hg/hour
$\lambda$	30 flushes/hour
$P_0$	18 in-Hg
$P_{low}$	14 in-Hg
$P_{lower}$	12 in-Hg

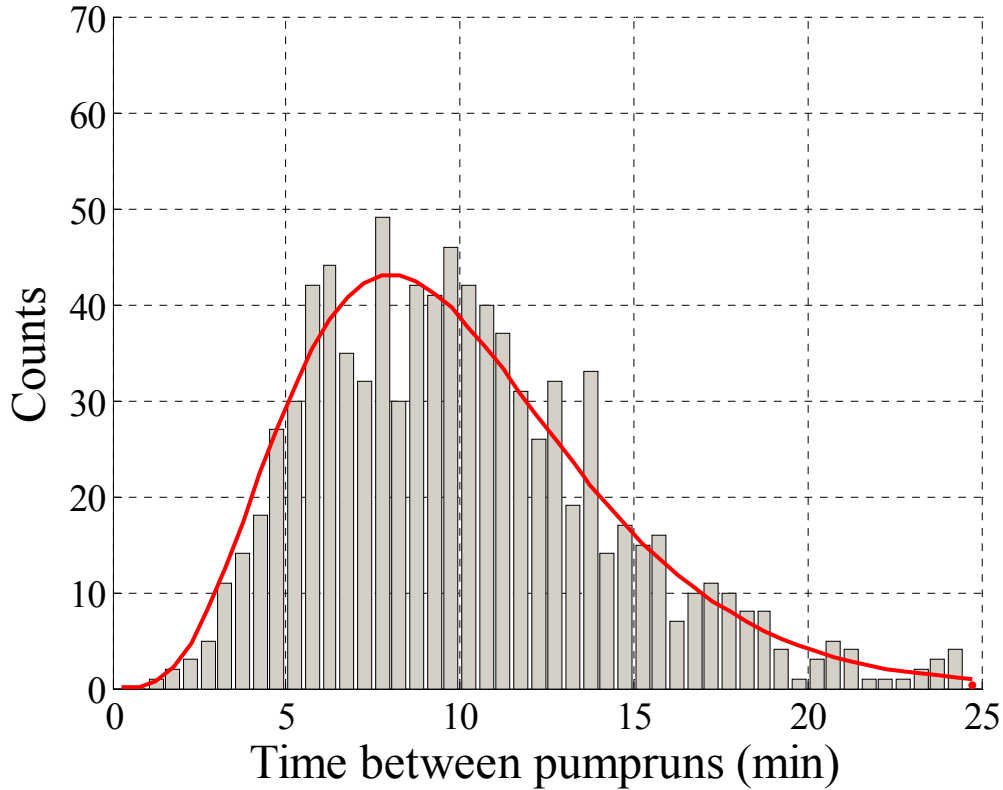
For the first demonstration, a flush size of 1.2 in-Hg/flush will be used. From equation (3.4),  $N_{max} = 4$  meaning that four flushes are required before the vacuum pumps energize. Based on the previous discussion of the Erlang PDF and on Figure 3-1, if  $N_{max} - 1$  flushes have occurred,  $P_{low}$  will never be reached if there is no leak in the system. It must be assumed that  $N_{max}$  flushes have occurred and thus  $k=N_{max}$  for the Erlang PDF. Figure 3-4 below shows the results with an Erlang PDF of order four ( $k=4$ ,  $\lambda=30$  flushes/hour) overlaid on the histogram.





**Figure 3-4: One week simulation with no leak,  $\lambda=30$  and Erlang of order 4 overlaid.**

Running the simulation again with a flush size of 0.9 in-Hg/flush. This time,  $N_{\max}=5$  and the results are plotted in Figure 3-5 on the same scale as the previous plot in order to see the shape differences.



**Figure 3-5: One week simulation with no leak,  $\lambda = 30$  and Erlang of order 5 overlaid.**

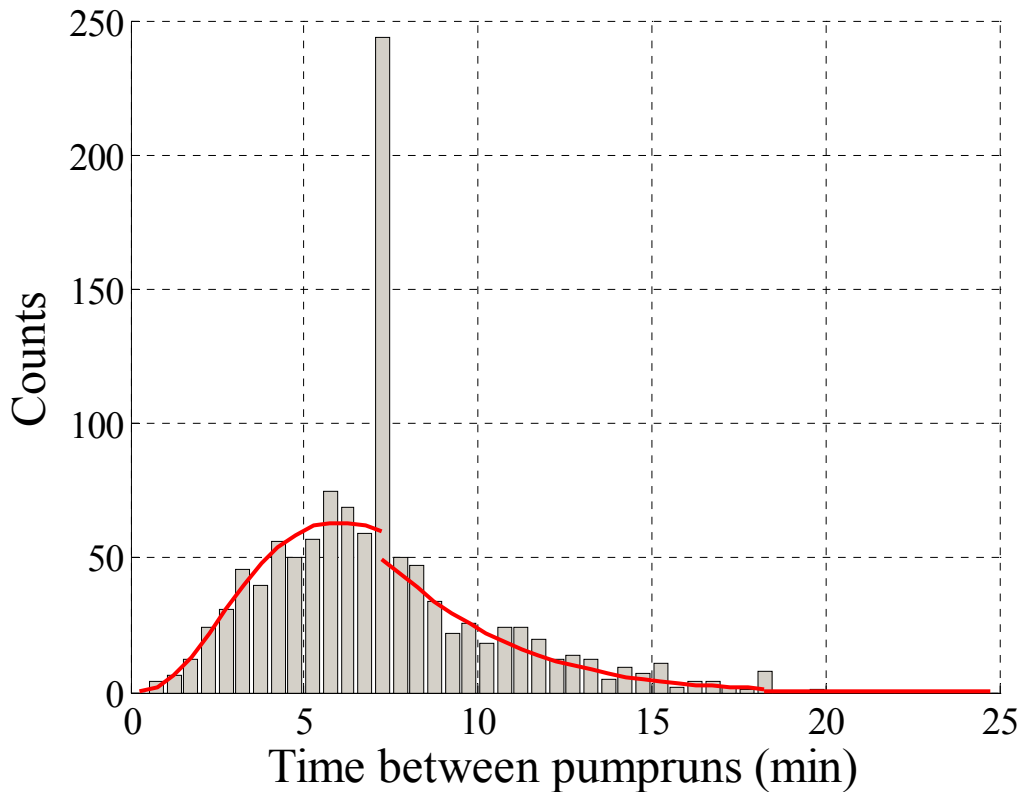
The phenomenon demonstrated in the lower plots of Figure 3-3 and which arises in the presence of a leak in the base model can be explained in a similar manner as the above. The height and the location of the “spikes” in the plots can be predicted given the size of the flushes, the high and low pressure setpoints, and the leak rate.

Using Figure 3-2 for a visual reference, it can be seen how the Erlang PDF relates to a resultant plot of times between pump runs. Based on the Figure 3-2, at time progresses from  $t = 0$  minutes up until  $t_3$ , there are four flushes required to start the pump. From  $t_3$  to  $t_2$ , there are three flushes required; from  $t_2$  to  $t_1$ , there are two flushes required; from  $t_1$  to  $t_0$ , there is one flush required; and at  $t_0$ , the pump is guaranteed to start without any flushes occurring. For each time period, as with the case of no leak, the order of the Erlang associated with that time period corresponds to the number of flushes required to start the vacuum pump. Thus, the orders of the Erlang would go from four to one respectively as each  $t_i$  is passed. As the order of the Erlang changes, there exists a discontinuity and is manifested as a “spike” in the histogram.

To demonstrate this idea, suppose that a 6 in-Hg/hour leak exists in the same system where the flush size is 1.1 in-Hg/flush. The calculated  $t_i$ 's are included in Table 3-2 and Figure 3-6 shows the histogram of times between pump runs with the corresponding Erlang PDF's overlaid. Note that the spikes are located at the calculated  $t_i$ 's that are within the range of the data. This is expected since the  $t_i$ 's indicate when a pump will energize.

**Table 3-2: Calculated "spike" times for a 6 in-Hg/hr leak and flush size of 1.1 in-Hg/flush.**

$t_i$	time (min)
$t_0$	40
$t_1$	29
$t_2$	18
$t_3$	7

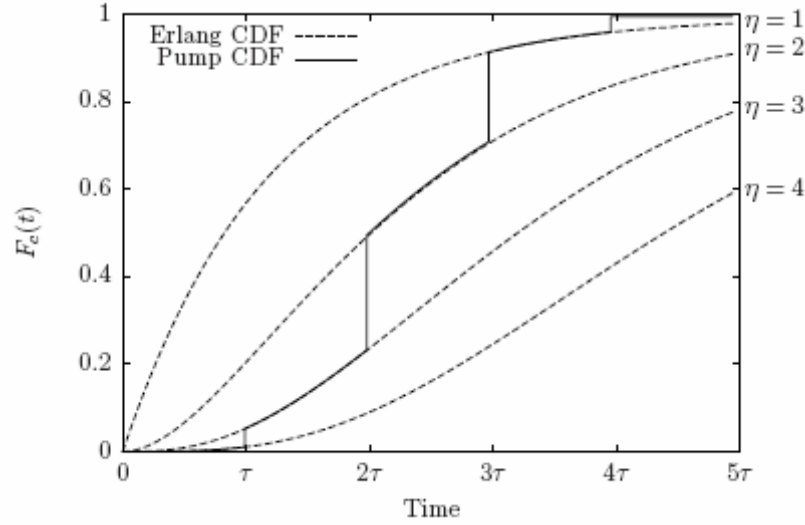


**Figure 3-6: One week of simulated data with 6 in-Hg/hour leak showing change in Erlang order.**

The spike height can be predicted based on this model as well. Once normalized by dividing by the total number of runs, the histogram shape still represents a PDF, so the integral under the entire shape is equal to 1. Since the Erlang orders change at the spike locations, the height of the spike must make up for the difference between the integrals of

the Erlang curves up to that point. This corresponds to the differences between the cumulative distribution functions at the  $t_i$ 's. For a more detailed description see ref [14]. Equation (3.7) is a word expression showing the spike height as a function of cumulative distribution functions ( $cdf_x$  = Erlang cumulative distribution function for  $k=x$ ) and Figure 3-7 shows the concept pictorially [14].

$$height_i = cdf_i - cdf_{i+1} \quad \text{for } i=1,2,\dots,N_{\max} \quad (3.7)$$



**Figure 3-7: Expected spike heights calculated from Erlang cumulative distribution function for spikes located at various times. Taken from ref [14] (η=Erlang order).**

Thus, using  $f_{erl}(k, \lambda, t)$  from equation (3.6) and the definition of the cumulative distribution function, the relation in equation (3.8) can be used to determine the expected height of the  $i^{\text{th}}$  spike located at time  $t_i$ . The result must be multiplied by the normalization factor in order to plot it on the same plot as the rest of the histogram.

$$height_i = \int_0^{t_i} (f_{erl}(i, \lambda, t) - f_{erl}(i+1, \lambda, t)) dt \quad \text{for } i=1,2,\dots,N_{\max} \quad (3.8)$$

## 4 Real System Modeling

### 4.1 Real System Characteristics

Formation of a basic model, as done in Chapter Three, is necessary to understand the underlying characteristics of the cycling system. The basic relationships must be understood before real world influences can be inserted into the model.

There are two primary sources of variation for the sewage system onboard the Seneca. First, variations in the system exist due to physics and due to mechanical aspects of the system components. Second, there is human variation in the system usage.

In order to more accurately reflect the real system, modifications had to be made to the base model simulation. The following are the modifications made and their effects will be further explored.

- The setpoints in the Seneca sewage system are not as described in the system manual
- The system is not perfectly sealed and has some small persistent leak in all conditions.
- The leak rate is not always constant. As the vacuum drops in the system, the leak rate lessens.
- The pressure drop per flush is not constant. Not only does the system pressure affect the drop, but each toilet or urinal has a different flush time which causes variation.
- The usage rate,  $\lambda$ , is not constant

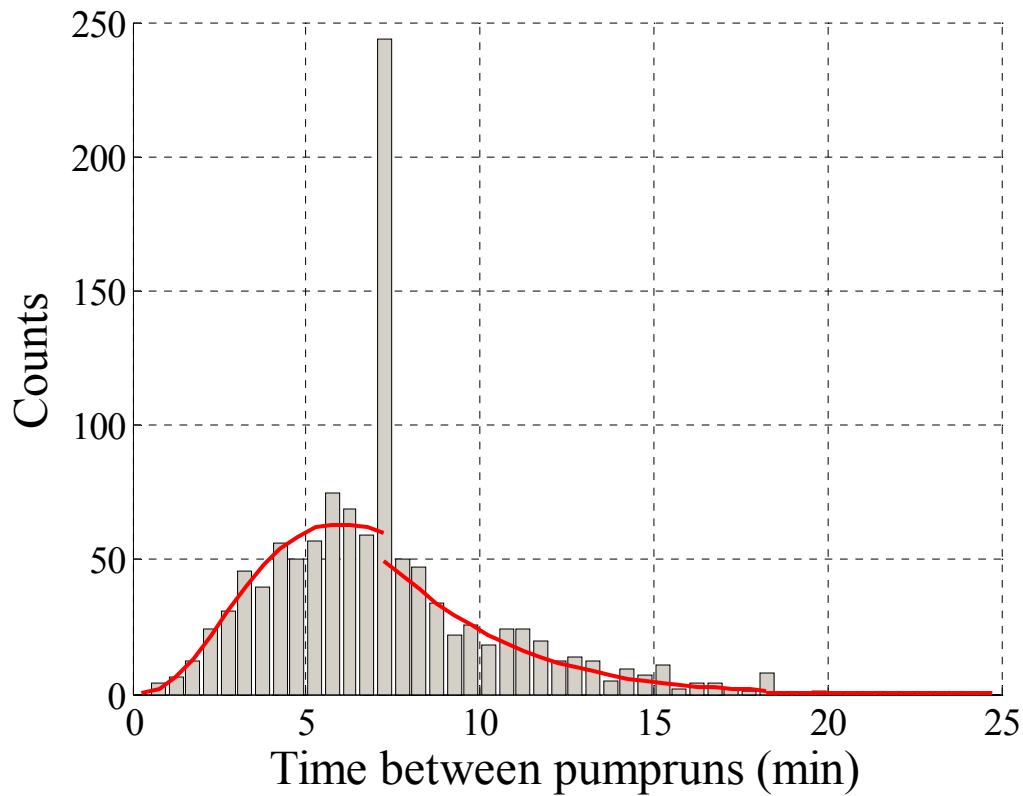
The first four modifications are the result of system variation and settings. The last modification is required due to the human aspect.

The pressure setpoints were not exactly the same as the factory settings, or as described in the system manual. The three pressure setpoints were originally set at 12, 14 and 18 in-Hg for the lower pressure second pump start, low pressure pump start and high pressure pump shut off respectively [15]. Using a separate pressure gauge, the setpoints were measured at 12.5, 13.5 and 17.5 in-Hg. Although these numbers do not vary greatly from the base model system, they are necessary to accurately model the system.

The system is not perfectly sealed and has a small persistent leak during all conditions. Although this leak can be accounted for in the leak added to the model, a

small constant leak rate was added to the model. The persistent leak was set to zero for the following demonstrations in order to not obscure the results.

To demonstrate the effects of each of the above variances, the random number generators in Matlab and in the Simulink model were seeded to a constant value. This allowed for a direct comparison of a histogram with no variation and a histogram with variation. Figure 4-1 below shows the histogram for a one week simulation with the realistic pressure setpoints,  $\lambda = 30$  flushes/hour,  $\Delta P_f = 1.1$  in-Hg/flush and maximum leak rate = 6 in-Hg/hour. Note the locations of the spikes are at 7 min and 18 min. Although the spike height at 18 min is low, there is evidence of a spike in that location.



**Figure 4-1: One week simulation baseline with 6 in-Hg/hour leak.**

## 4.2 Leak Rate Variation

The next variation arises from the fact that any leak rate is not constant, since the rate depends on the differential pressure between the system and the atmosphere. In order to include this effect in the simulation, a model of the system pressure was needed. The pressure in the system was modeled according to the first order differential equation

$$\frac{dP}{dt} = -cP, \quad (4.1)$$

whose solution is

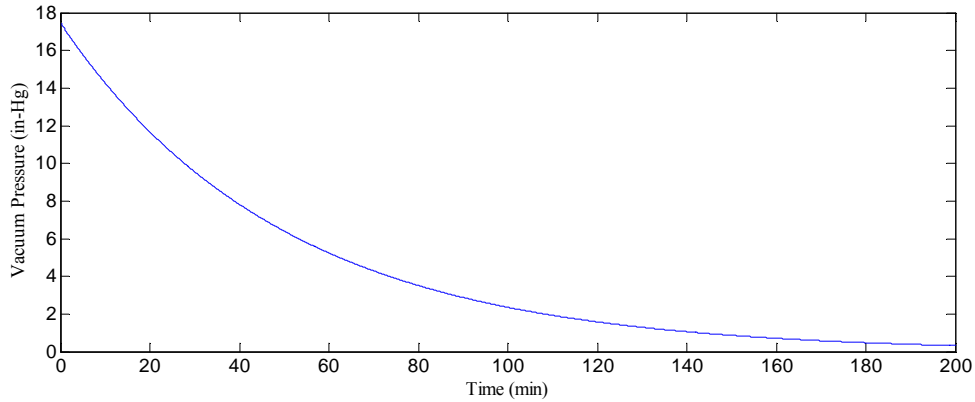
$$P(t) = P_0 e^{-ct}. \quad (4.2)$$

where  $P_0$  is the initial system pressure. The  $c$  value depends on the leak size and can be approximated using equation (4.1) at  $t=0$  by

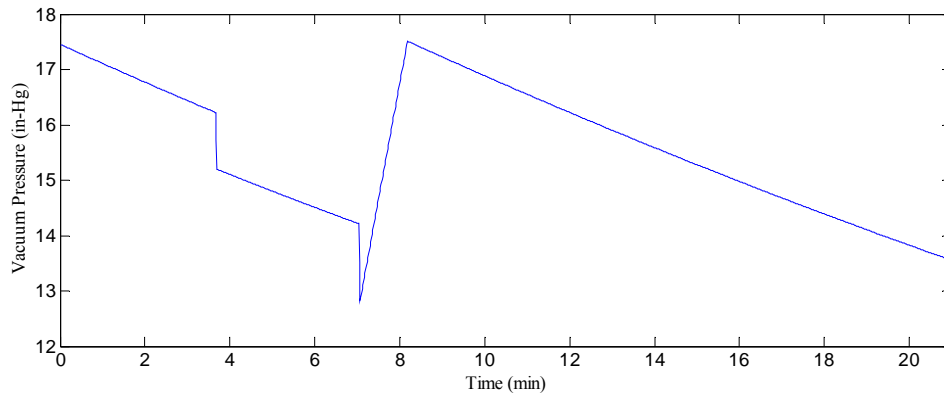
$$c = -\frac{\frac{dP(0)}{dt}}{P(0)} = -\frac{Maxleakrate}{P_0}, \quad (4.3)$$

where *Maxleakrate* is the highest leak rate obtained when a leak is inserted into the system while the system pressure is at the high pressure setpoint.

In the time domain, if a leak were installed and the vacuum pumps did not recharge the system, the pressure would drop off rapidly at first, and as the pressure differential lessened, so would the rate at which the pressure changed. The simulation was modified to account for this and Figure 4-2 shows the simulated system pressure in a case where the vacuum pumps did not energize to raise the pressure and Figure 4-3 shows the simulated pressure trace for normal operation of the sewage system. The results in Figure 4-3 can be compared to the actual pressure traces in Figure 2-7.



**Figure 4-2: Simulated system pressure over time given that the vacuum pumps to not energize to raise pressure.**

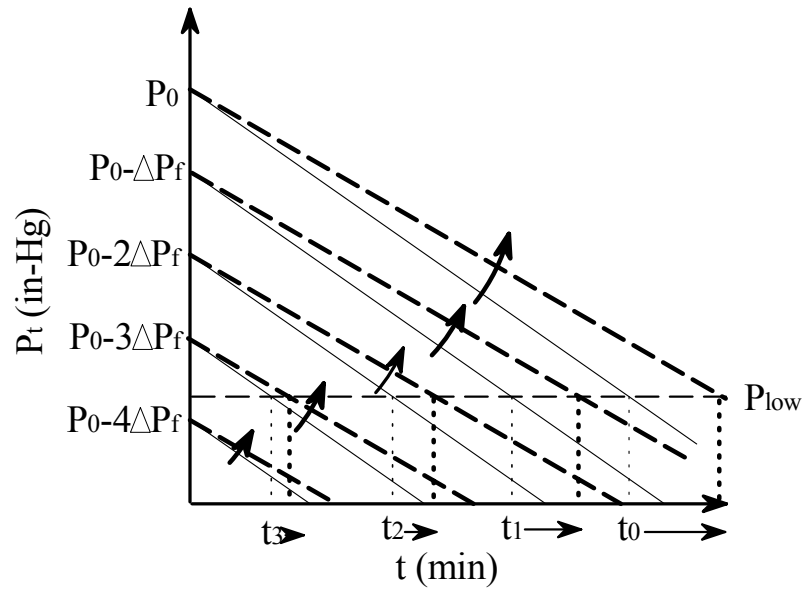


**Figure 4-3: Simulated pressure trace in normal operating range of vacuum system.**

Based on the model, given the high and low pressure setpoints in the Seneca system, the leak rate at the low pressure setpoint should be approximately 77% (=13.5 in-Hg/17.5 in-Hg) of the *Maxleakrate*. To confirm the model, gas flow meters were installed in the sewage system. A smaller 150 SCCM flow meter and a larger 100 SCFH flow meter both showed that the variation in flow did depend on the pressure in the system with the air flow at the lowest pressure approximately 75-78 % of the air flow at the highest pressure thus showing that the pressure model is adequate.

The maximum leak rate was the only leak rate used in the base model whereas the simulation adjusts the leak rate according to equation (4.1). The effect on the spike times can be shown in Figure 4-4. The lower leak rates translate to a lesser slope on the pressure lines and the spike times shift to the right as shown. The amount of the time shift is dependent on how much the leak rate changes. The leak rate is now a range of leak rates dependent on system pressure, so the slope change and subsequent spike time shift is a distribution vice a singular number. Instead of a tall narrow spike at one time, the resultant distribution of spike times manifests itself in a wider, shorter spike centered on a new time. The center of the spike distribution can be estimated using the expected value of the leak rates. Since the leak rate model is linear and the system pressures are all equally likely, the expected value of the leak rate is simply the average of the leak rates. For the Seneca system where the leak rate ranges from 77% to 100% of the maximum leak rate, the expected value is thus 88.5% of the maximum leak rate. Note that the amount of the spike time shift is not the same for all the times. The spike associated with  $t_0$  moves the farthest whereas the spike associated with  $t_{N_{max}-1}$  moves the least.



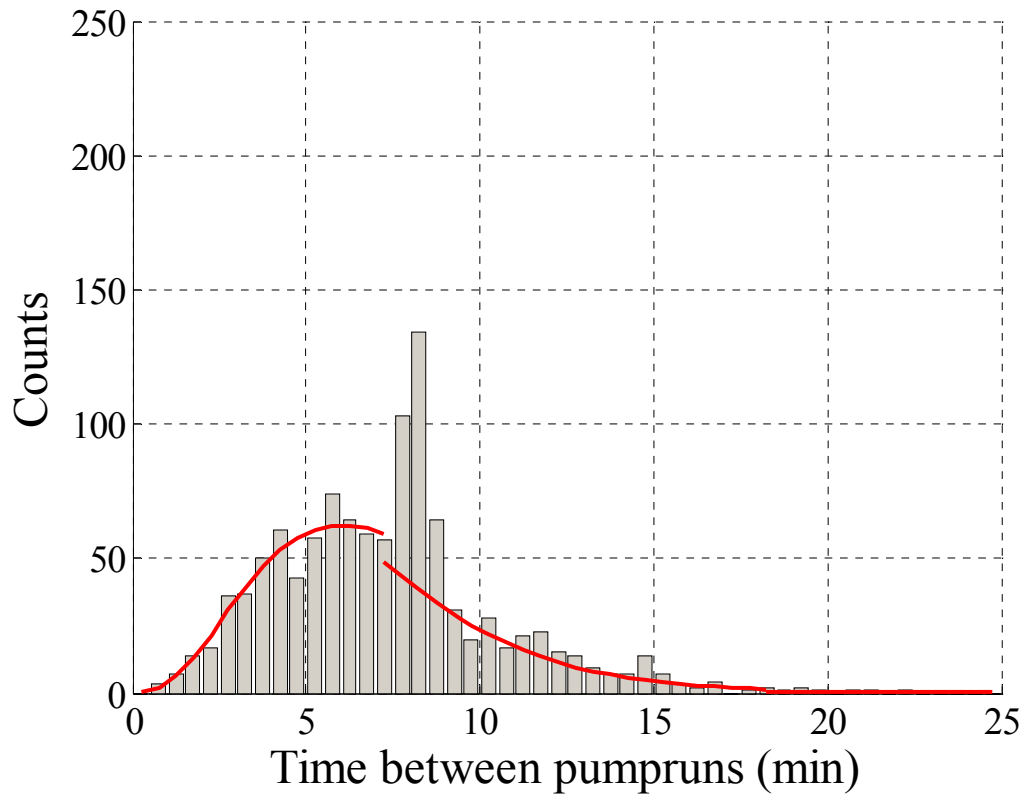


**Figure 4-4: Effect of varying leak rates on spike times.**

The predicted spike locations are included in Table 3-2. Using the expected value of the leak rate to be 88.5% of the max leak rate and Figure 4-4, the new times can be calculated and are included in Table 4-1. These times are the expected center of the new spikes with a small distribution of the spike on either side. The resulting histogram using the same random number “seeds” as the baseline in Figure 4-1 is shown in Figure 4-5 with an ideal Erlang fitted curve overlaid.

**Table 4-1: Expected spike times with variable leak**

$t_i$	time (min)
$t_0$	45.2
$t_1$	32.8
$t_2$	20.33
$t_3$	7.9



**Figure 4-5: One week simulation with leak rate variation**

As predicted, the spike shifts to the right and decreases in height due to the distribution of leak rates. The location of the spike again lines up with the calculated  $t_i$ 's. The effect of a varying leak rate that is a function of system pressure is thus shown to have a “smoothing” effect on the histogram of time between pump runs.

### 4.3 Flush size variation

The next variation that is considered is the variation in the size of the flush or any system usage event. Since no two toilets or urinals are exactly the same, the flush size cannot be assumed constant as is done in the base model. Also, the duration of a usage event is not instantaneous as was assumed in the base model.

The system pressure also affects the size of a usage event because the changes in differential pressure between the system and the atmosphere cause the flush sizes to vary.

This effect was not singled out and simulated but assumed to be taken into consideration with the distribution of flush sizes around a mean size.

To verify the pressure drop associated with a system usage event, a pressure sensor was installed in the system and data recorded alongside the usual NILM data. An example of the pressure trace with an installed leak and the vacuum pump power associated with the pump runs is shown above in Figure 2-7.

The results of testing showed that there were predominantly two usage event sizes, approximately 0.80 in-Hg/flush and 1.30 in-Hg/flush with a small amount of variation around each of those levels and those flushes typically last approximately two seconds. The duration of the usage event was not analyzed separately but was included in the Simulink model in order to better model the real system.

To begin the analysis, it was assumed that there was only one flush size with a distribution around that flush size. The effect of variation in the size of a usage event is demonstrated below in Figure 4-6. For each  $N$  value, the variation creates a distribution of possible pressures. For the  $t_i$ 's, the result is a distribution of times on either side of the base model  $t_i$ . Again, this distribution manifests itself as a wider, shorter spike on the histogram. The resulting histogram for the baseline case with a flush size uniformly distributed between 1.0 and 1.2 in-Hg/flush is shown in Figure 4-7, again with the base model Erlang distributions overlaid. Any previous variations examined were removed in order to show the singular effect of flush size variation.

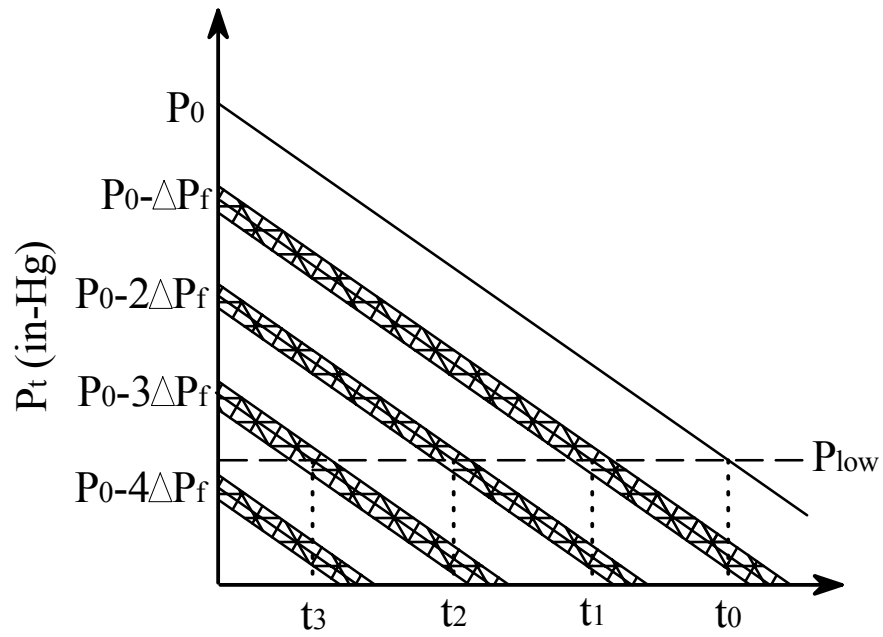


Figure 4-6: Effect of varying vacuum loss due to a usage event.

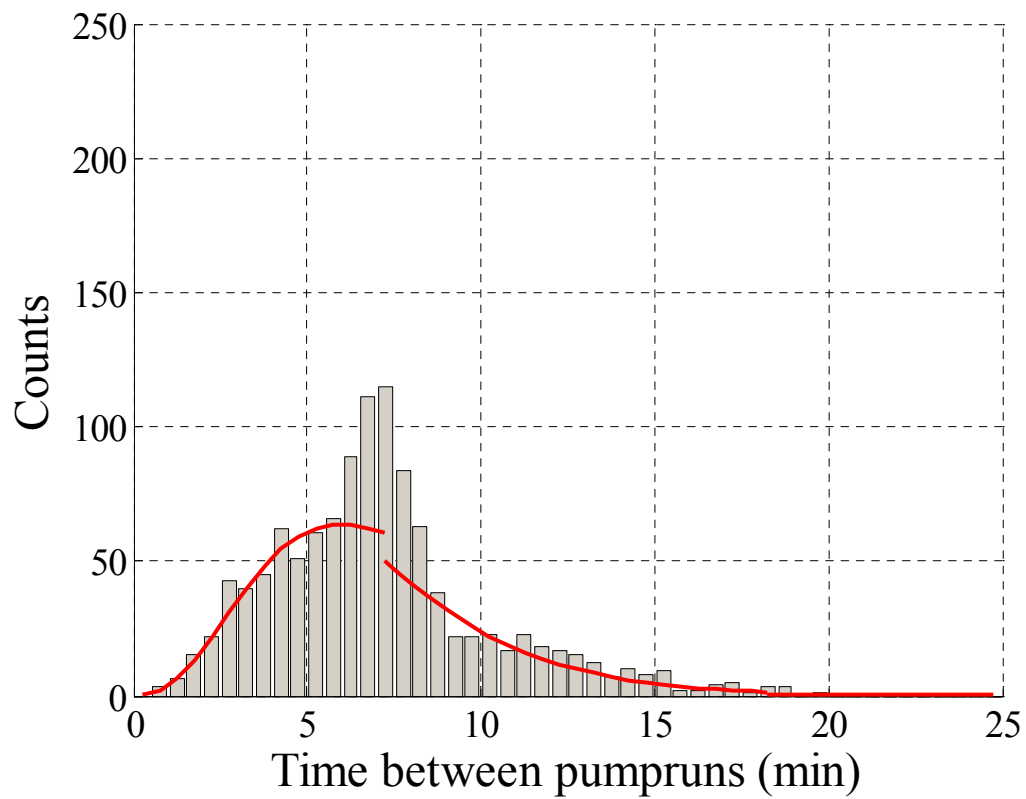


Figure 4-7: One week simulation with usage event size variation.

As predicted, the spike height decreased and created a distribution about the expected  $t_i$ . Next, consideration was given for two predominant sizes of usage events. The change on Figure 4-6 would be that there would be two different  $\Delta P_f$ 's which would result in twice as many  $t_i$  values. The resulting histogram would have multiple spike locations with a distribution around each  $t_i$  similar to what is seen in Figure 4-7.

#### 4.4 System Usage Rate Variation

The variation that has a tremendous effect on the time between pump runs is the rate of usage on the system. Variation onboard a ship, either at sea or inport, is very difficult to quantify. The base model simulation used the same  $\lambda$  for all times of the day.

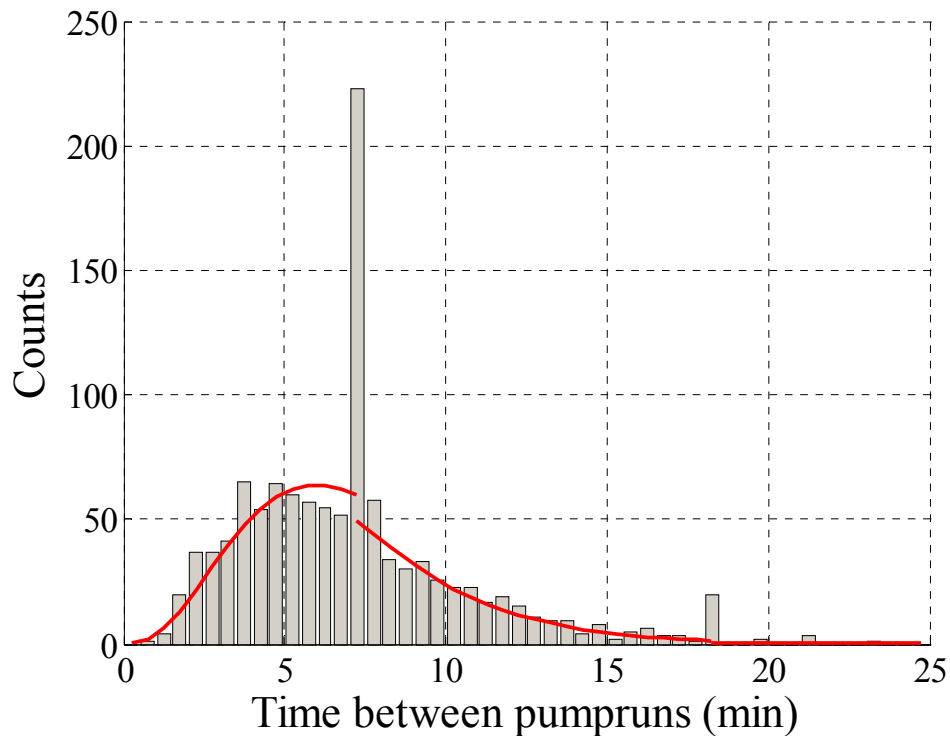
Ideally, a continuously varying usage rate could be determined and used in the model, but determining the precise rate would be very difficult. Analysis of the data from Seneca indicated that there tended to be three distinct time periods during the day which had different usage rates. The times corresponded well to the work day either at sea or inport. Three eight hour time segments were chosen ranging from 0600 to 1400 ("work hours"), 1400 to 2200 ("evening") and 2200 to 0600 ("nighttime"). The usage rates were lowest during the nighttime time while a majority of the crew sleeps. The work hours and evening time frames appear to have similar usage rates, although the evening is usually slightly higher. This is expected since the crew usually has more free time in the evening and is not consumed with on-watch activities and don't have time to use the restrooms.

The usage rate for each eight hour time period is essentially the time weighted average of the usage rates during that time. Since  $\lambda$  is a function of time, the system usage process is referred to as a non-homogeneous Poisson process. The nonlinear time transformation shown in equation (4.4) can reduce the problem to a homogeneous Poisson process [13]. Although unable to determine the exact  $\lambda(t)$  throughout each time period, it was possible to estimate the usage rates based on the number of pump runs in the period and the average sizes of the flushes.

$$\lambda_{avg} = \int_0^t \lambda(s) ds \quad (4.4)$$

It is important to note that although the pressure is “reset” by the vacuum pumps, the flushing is independent of the pumping. This means that the pumps deenergizing when the pressure reaches  $P_0$  is not a true renewal event as is ideal for a Poisson process. Even though a flush can occur a few seconds before the vacuum pump secures, the base model considers  $t = 0$  when the pump secures. This means that the time to the first flush after a pump securing in the base model is exponentially distributed from  $t=0$  even though it should be distributed from the last flush. In reality, and in the Simulink model, the flush times are completely independent of the pump cycles and thus the measured time between pump stops and pump starts is reflected in the time between pump runs histograms.

Using three different values of  $\lambda$  (20, 34, 36 for the nighttime, work hours and evening respectively —with a mean of 30 as was used for the all-day rate in the base model) in the baseline model results in the histogram shown in Figure 4-8. It can be seen that the lower  $\lambda=20$  value tends to “fill out” the right side of the distribution as evidenced by the taller spike at 18 minutes and the appearance of times greater than 20 minutes. The two other higher  $\lambda$  values tend to “fill out” the left side of the distribution as evidenced by the slightly taller bins in the 1-3 minute range.



**Figure 4-8: One week simulation with three different lambda values.**

Not only does the lambda vary throughout the day, but it also varies from day to day. The activities of the crew, the missions being performed, the food served, and the overall health of the crew tend to vary the lambda values from day to day. The model was altered to account for this variation and is included in the “prep” file when determining the flush times.

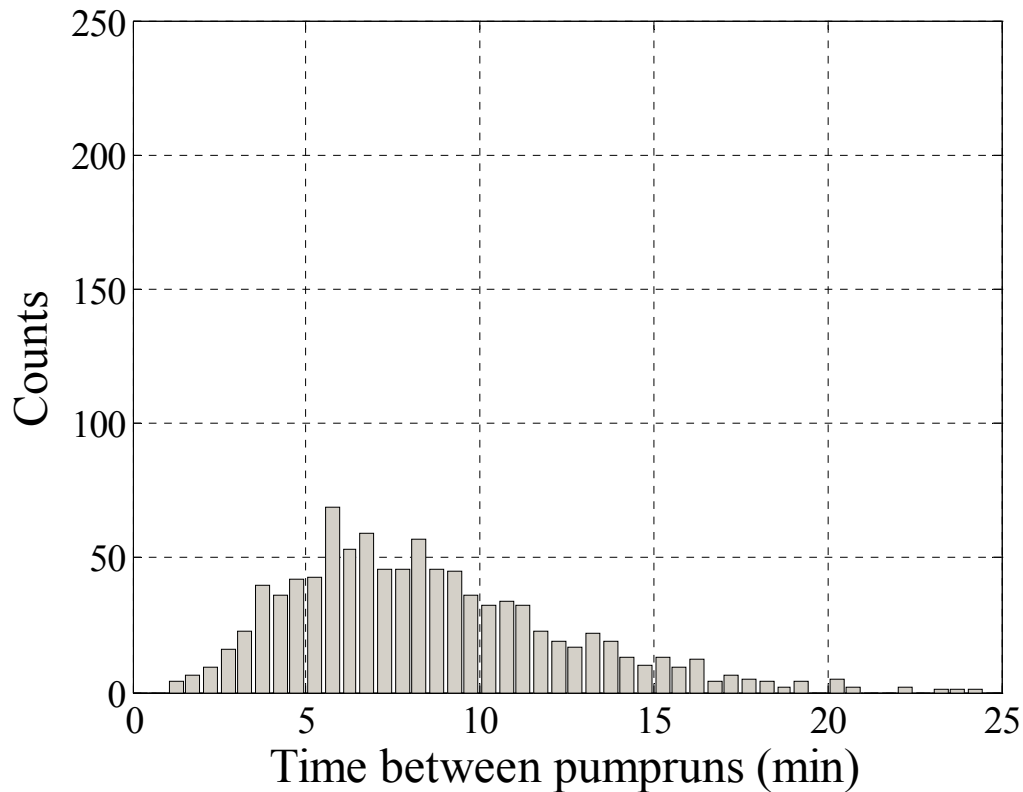
## 4.5 Compilation of Variation in All Factors

So far each of the characteristics of the system that can have variation has been analyzed individually. In reality, they all can vary together and change the shape of the histogram. Table 4-2 below lists the allowed variation of each parameter and Figure 4-9 shows the compilation of all the effects of all the variations on the histogram of times between pump runs.

**Table 4-2: Parameter variation allowed in the simulation model.**

<u>Parameter</u>	<u>Variation Input into Simulation</u>
Leak Rate	Variation linearly dependent on system pressure
$\Delta P_f$	Flush sizes uniformly distributed between at 0.6-0.72 and 1.0-1.2.
$\lambda$	nighttime = 20 flushes/hour work hours = 34 flushes/hour evening = 36 flushes/hour Note: each allowed to vary $\pm 20\%$ during the 8 hour period

As witnessed in Figure 4-9, the smoothing effect of all the variation makes the presence of a leak not as obvious as in Figure 4-1. Diagnosis of the leak thus becomes more complicated and determination of the leak size is even more difficult. Chapter Five investigates the possible diagnostic indicators and the best method of determining the size of leak in the system.

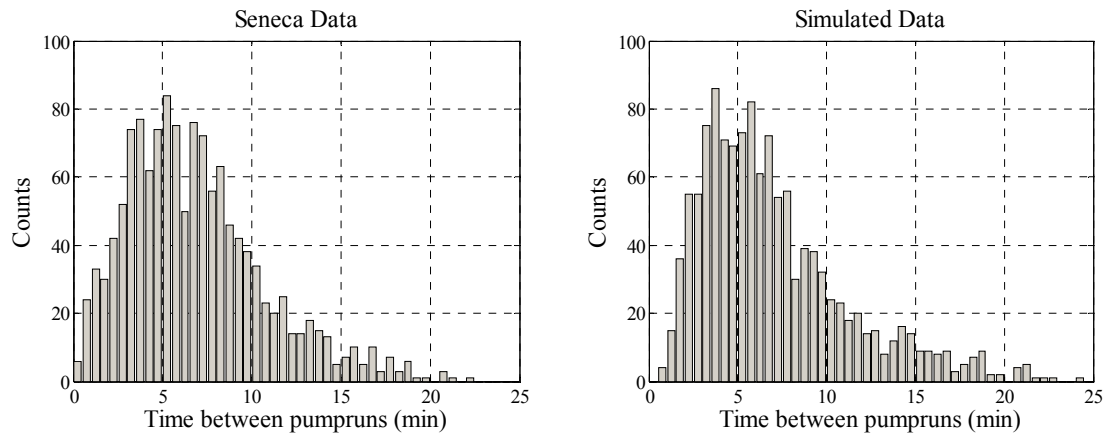


**Figure 4-9: One week simulation with compilation of parameter variations.**

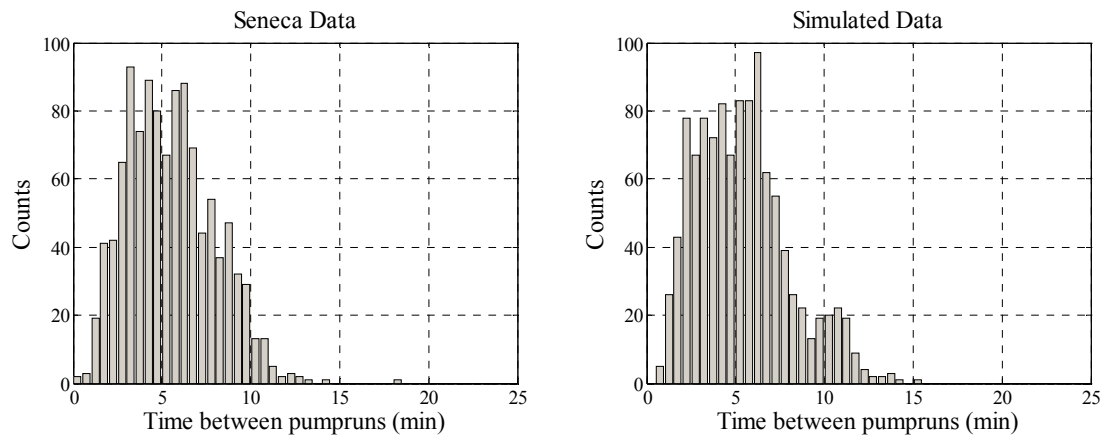
## 4.6 Simulation Results

Putting all the variations and adjustments into the model, it can now be tested against actual data. The figures below show comparisons of actual Seneca data and simulated data for the same time periods. The two predominant flush sizes were simulated to match what was seen on the ship as well as the duration of usage events. Variation as discussed in the previous sections was incorporated and adjusted to match real variation as closely as possible. The mean daily usage rate for used in the simulation was 30 flushes per hour for underway simulations. The comparisons are of a system with no leak, a system with a 12 in-Hg/hour leak and a system with failure of the check valves between the tank and vacuum pumps.

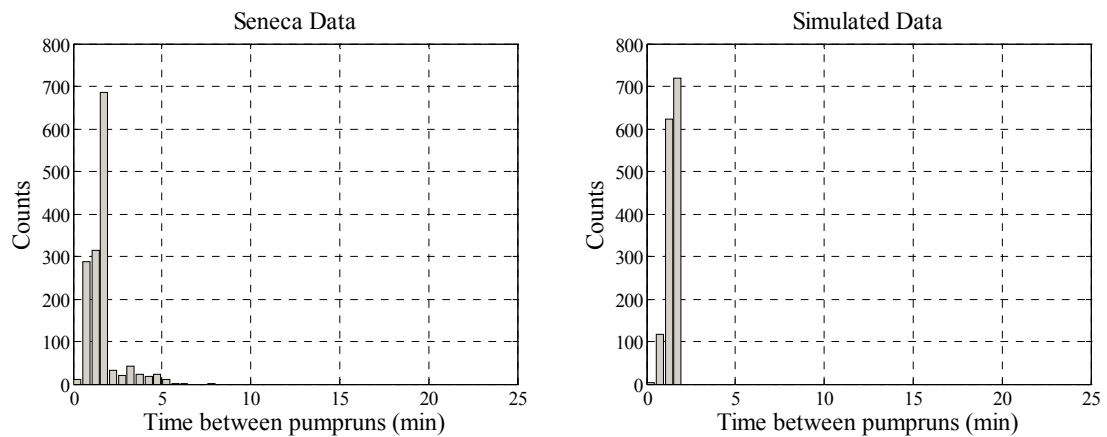




**Figure 4-10: Comparison of Seneca underway data and simulated underway data for a seven day period with no leak. Seneca total number of runs = 1297, simulated total runs = 1288.**



**Figure 4-11: Comparison of Seneca data and simulated data for a five day period with 12 in-Hg/hr leak. Seneca total number of runs = 1102, simulated total runs = 1100.**



**Figure 4-12: Comparison of Seneca data and simulated data for a three day period with check valve failure. Seneca total number of runs = 1476, simulated total runs = 1463.**

As can be seen by the histograms, the simulated data varies slightly from the Seneca data. The simulated distributions are slightly narrower, but the numbers of runs are very near to each other as well as the mean times between runs. The amount of variation, especially in usage rates, onboard the ship is difficult to simulate, but the Simulink and Matlab model adequately replicate the times between pump runs onboard the ship.

## 5 Diagnostic Indicator

Diagnostic software for this vacuum assisted system must be able to determine its overall health. The primary concern is the presence of a leak in the system because a leak that is not caught and fixed can cause excessive wear on the vacuum pumps and wastes electrical energy. Leak detection is difficult, however, as an elevated usage rate from such instances as a sick crew or the addition of a large group of people onboard can also cause a change in the histogram. The goal of the diagnostic method is to determine if the usage of the system has changed and if that change was caused by a leak.

When comparing leak versus no leak data, either from the ship or from simulation, there are a number of indicators of change. Although a visual inspection of the histogram of times between pump runs is one way to determine the presence of a leak, the most convenient diagnostic tool is one that performs the detection process automatically without any human intervention.

### 5.1 Possible Diagnostic Methods

There were several quantitative methods used to analyze the results, both from the ship and from simulation. Each of the methods has its own strengths and weaknesses, and these are explored in the following sections.

Since changes in system operation are reflected in how often the pumps operate, the first proposed method analyzes the mean time between runs and the total number of runs over a given period. Another diagnostic method is the detection of discontinuities in the histogram. A third method involves trending the parameters of a curve fitted to the histogram data. Lastly, an analysis of the time each pump is energized is presented.

#### *5.1.1 Mean Shift Test and Total Number of Pump Runs Test*

The mean time between pump runs and the total number of pump runs are fairly strong indicators of a change in the system, but they do not discern between high usage and the presence of a leak. Table 5-1 below contains sample mean times between pump runs and the total number of runs for various conditions, both actual and simulated. Note

the decrease in mean times and the increase in number or runs in the case of a leak and in the case of increased usage.

**Table 5-1: Samples of means and total number or pumps runs.**

<u>Seneca Underway Data</u>	<u>Mean time (min)</u>	<u>Three day total</u>
No leak – August 2005	6.71	555
No leak – December 2005	7.08	519
12 in-Hg/hr leak – November 2005	5.39	682
12 in-Hg/hr leak – January 2006	5.28	687
<u>Simulation Data</u>		
No leak (run1)	6.74	554
No leak (run 2)	6.72	555
12 in-Hg/hr leak (run 1)	5.29	674
12 in-Hg/hr leak (run 2)	5.33	668
10 flush/hour increase (run 1)	5.29	675
10 flush/hour increase(run 2)	5.15	692

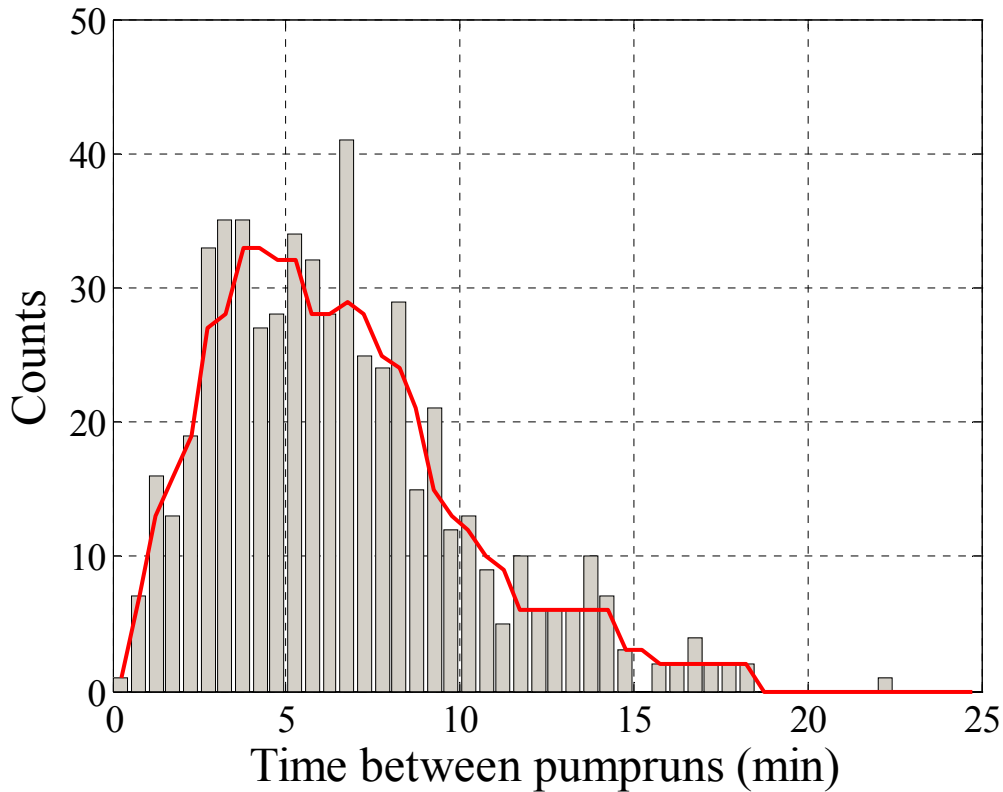
Although the mean shift test does not give any indication of what is causing the change in system behavior, it is a definite indication of some change in system operation. Perhaps a better indicator of system operation change is the total number of runs. Just like the mean time between runs, though, the total number of runs cannot discriminate between leaks and increased usage rates. The total number of runs can be used as an initial indicator, but another test must be used to determine if a leak exists in the system or if the usage rates have increased.

### *5.1.2 Discontinuity Detection Test*

One distinct difference between leaks and usage rate changes is that leaks, especially large ones, distort the expected histogram. The result is the formation of spikes and sharp edges. A way to find such discontinuities is to use a median filter on the binned times. A median filter finds the median value of the data on either side of the current bin along with the data in that bin. For example, a median filter with a window size of seven examines the three bins on either side of the current bin along with the center bin to find a filtered value [21]. In equation form, for a filter of size  $(2*N+1)$  applied to bin data,  $y(t_i)$ , the median filtered value for each time,  $y_{\text{filt}}(t_i)$ , can be expressed as

$$y_{filt}(t_i) = median[y(t_{i-N}), \dots, y(t_{i-1}), y(t_i), y(t_{i+1}), \dots, y(t_{i+N})]. \quad (5.1)$$

When applied to a histogram of time between pump run data, the result is a smoothing of the bin counts. For instance, any abrupt changes in bin counts from a spike would be filtered out. Figure 5-1 shows a sample histogram for three days of underway, no leak Seneca data overlaid with the same data median filtered using a window size of seven. As can be seen, the thick line is more “smooth” and the large differences between bins are filtered out.

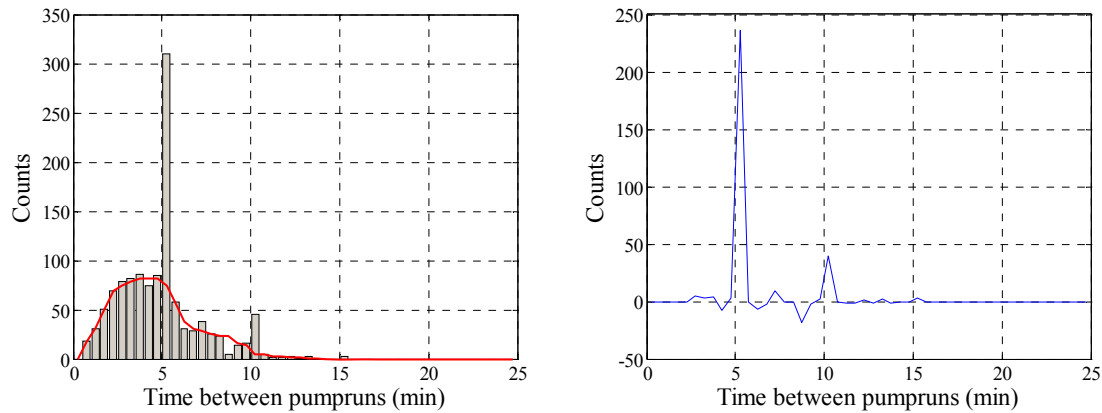


**Figure 5-1: Histogram for three days of underway, no leak Seneca data overlaid with median filtered data (window size = 7).**

#### 5.1.2.1 Base Model System Application

The median filter works very well with the base model. The lack of variation produces sharp spikes and large discontinuities that are easily captured using a median filter. An example of a data set with the median filtered results plotted over the data is shown in leftmost plot of Figure 5-2. By subtracting the filtered data from the original data, the presence of spike becomes evident and easily detected using a simple threshold

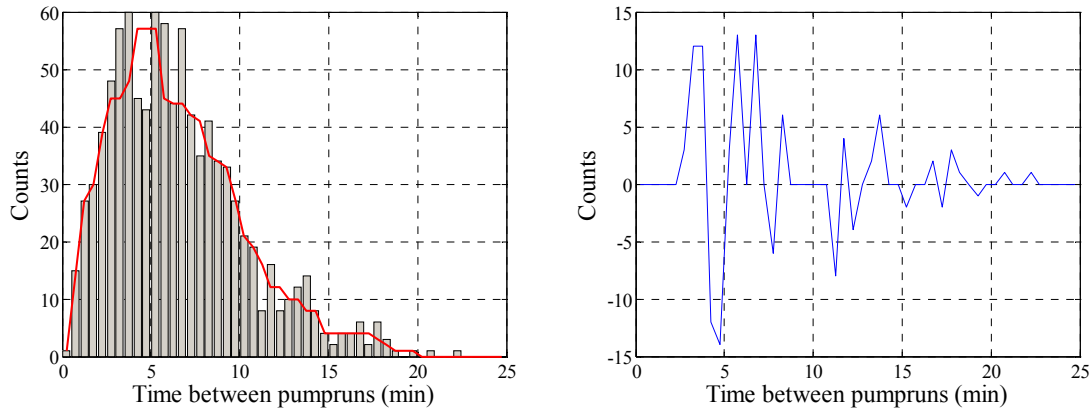
method. The rightmost plot of Figure 5-2 shows this result for the data plotted on the left. . Note the large values on the rightmost plot that correspond to the location of the large spikes on the left plot.



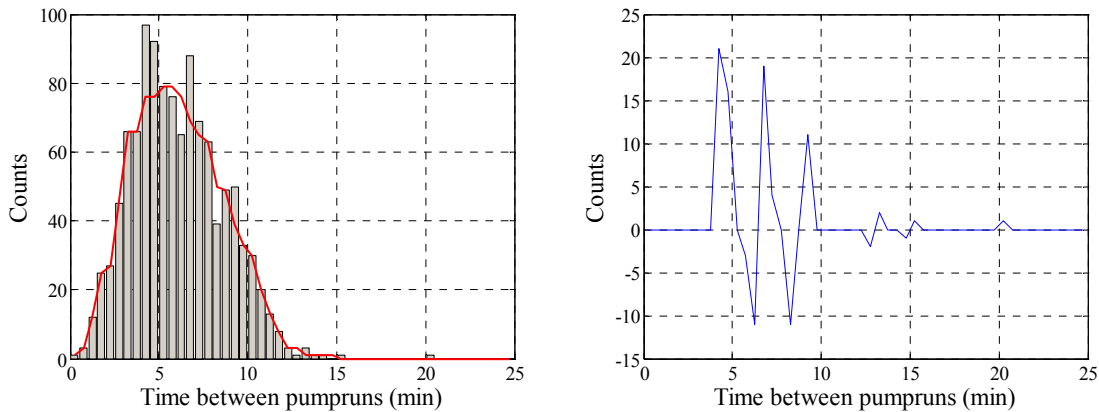
**Figure 5-2: Median filtering example on base model data. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data.**

#### 5.1.2.2 Real System Application

Using the same technique on real data is not as useful. Due to the variance in the shipboard system and in the usage rates, a large spike does not always exist as seen in the base model case. Figure 5-3 below shows median filtered data for five days of data with no leaks and Figure 5-4 shows the same comparison on five days of data with an installed leak. Note that the magnitude of the differences in the right plots is nearly the same in both cases. The maximum difference ranges between 10 and 20 whereas it was over 200 for the base model case. This problem prevents us from being able to use discontinuity detection to find small leaks.

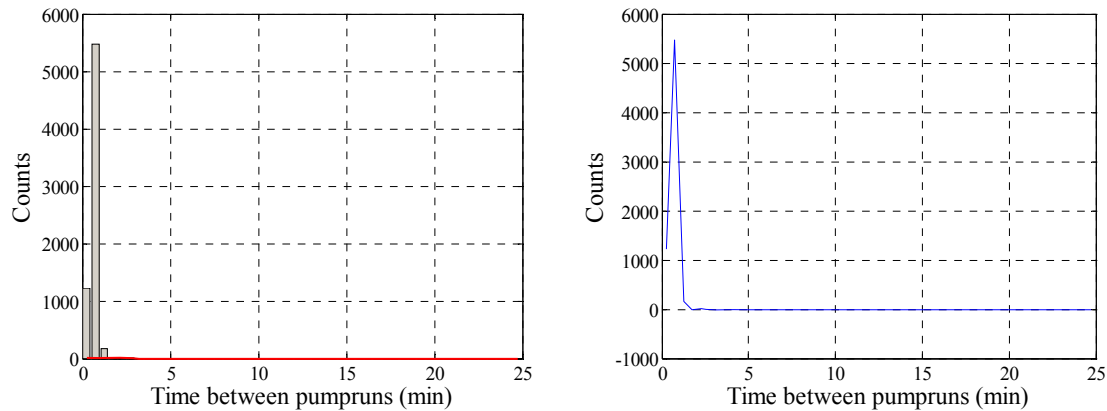


**Figure 5-3: Median filtering example on five days of Seneca data with no leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data.**



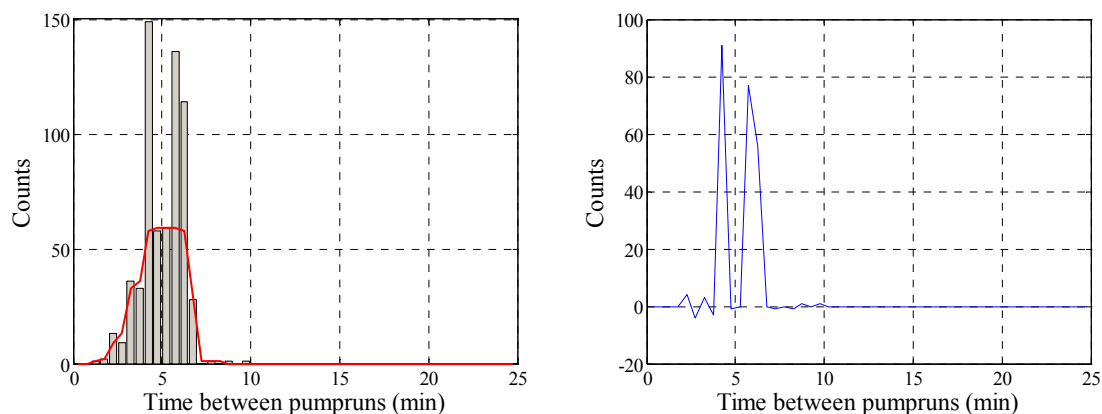
**Figure 5-4: Median filtering example on five days of Seneca data with leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data.**

Discontinuity detection is still a useful tool, especially when usage drops significantly or if there is a massive leak. For a massive leak in the system while underway or inport, as seen when the check valves located between the pressure tank and the vacuum pumps fail, the times between pump runs abruptly end after the first few bins. Figure 5-5 shows five days of Seneca data with check valve failure. Note that the median filter easily identifies the sharp difference between the second and third bins.



**Figure 5-5: Median filtering example on five days of Seneca data with check valve failure. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data.**

For inport data where the number of flushes and subsequent pumps runs is less, the presence of a large leak manifests itself in an abrupt end to the histogram. The sharp end roughly represents the time required for the system pressure to drop from the high pressure setpoint to the low pressure setpoint without any flushes occurring. If, for instance, the leak is approximately 40 in-Hg/hr, then the approximate time required for the system pressure to drop from the high pressure to low pressure setpoints would be six minutes. Therefore, an abrupt end to the histogram would be expected at approximately six minutes. Figure 5-6 shows inport data with a large leak and the median filtered difference.



**Figure 5-6: Median filtering example on three days of inport Seneca data with large leak. The left plot shows the histogram with the median filtered data overlaid and the right plot shows the difference between the original data and the median filtered data.**



### 5.1.3 Parameter Estimation and Trending

#### 5.1.3.1 Gamma Probability Density Function

Given the analysis results of the base model cycling system, one natural diagnostic method is to determine how closely the observed data fits to an Erlang PDF. Programs such as Matlab or Mathcad can do this easily and both were used. Recall from Chapter Four, however, that a real cycling system behaves slightly differently than the base model system described in Chapter Three. As a result, measured data will never truly follow an Erlang distribution. Based on numerous field observations, it was determined that a reasonable model for the actual distribution is the gamma PDF. This distribution is commonly encountered in reliability studies that aim to solve the similar problem of determining the distribution of times between equipment failures [23]. The gamma distribution is given by the equation:

$$f_{\text{gamma}}(k, \lambda, t) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{\Gamma(k)} \quad (5.2)$$

where the gamma function,  $\Gamma(k)$ , is defined as [22]

$$\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx. \quad (5.3)$$

Note that if  $k$  is a positive integer,

$$\Gamma(k) = (k-1)!, \quad (5.4)$$

and  $f_{\text{gamma}}(k, \lambda, t)$  thus reduces to the Erlang PDF presented in equation (3.6). This ability to describe the base model system behavior makes the gamma PDF intuitively pleasing.

In order to use the gamma model as a diagnostic tool, one must do more than simply estimate the parameters of the expected distribution. In particular, it is necessary to consider some sort of goodness-of-fit test or parameter trending that can indicate if the behavior of the system is beginning to deviate from its expected patterns. The remainder of this section contains brief discussions of the numerical methods used to estimate the parameters of the model in equation (5.2). Also included is a description of two diagnostic methods that rely on the results of this estimation.

#### 5.1.3.2 Non-Linear Least-Squares

One method for estimating the values of  $k$  and  $\lambda$  is to use non-linear least-squares, which selects parameter values that minimize the objective function

$$g(k, \lambda) = \sum_{bins} (O - E)^2, \quad (5.5)$$

where O is the observed bin count and E is the expected, or calculated, bin count [23]. Computer tools can easily and quickly perform least-squares fits using a method such as that of Levenberg and Marquardt [24].

Although least-squares curve fitting is a powerful analysis tool, it is not necessarily the best method to use when estimating the parameters of a density function. As expressed in the objective function in equation (5.5), least-squares requires that all of the measured data be placed into histogram bins. This procedure can be problematic, as any bins with few entries will fail to satisfy the requirements of Gaussian statistics, which is necessary when using the least-squares method [23]. As a result, other more general methods were considered and used, but for completeness, least-squares is considered in section 5.1.4.

#### 5.1.3.3 Maximum Likelihood Estimation Method

Another estimation technique considered in this thesis is the use of maximum likelihood estimators (MLE). In this approach, the data is not binned; rather, the model parameters are estimated using the individual time measurements.

Given the model  $f(x; \theta)$  and  $n$  observations of the random variable  $\mathbf{X}$ , the likelihood function of a random sample is defined as

$$L(x; \theta) = \prod_{i=1}^n f(x_i; \theta), \quad (5.6)$$

where  $\theta$  is the set of parameters that describe the underlying model. When the model is the gamma PDF,  $\theta$  includes the parameters  $k$  and  $\lambda$ .

As shown in [22], the maximum likelihood method estimates the parameters  $\theta$  by maximizing the likelihood function. When estimating the parameters of the gamma distribution, it is more convenient to perform the equivalent operation of maximizing the log of the likelihood function.

For the gamma distribution, the maximum likelihood equations are [22]:

$$\bar{x} = \frac{k}{\lambda} \quad (5.7)$$

$$\ln\left(\frac{1}{\lambda}\right) + \Psi(k) = \ln(G) \quad (5.8)$$

where  $G$  is the geometric mean of the sample  $x$  and  $\Psi(k)$  is the digamma function [22].

These two can be combined into

$$\ln(k) - \Psi(k) = \ln\left(\frac{\bar{x}}{G}\right) \quad (5.9)$$

Solving equation (5.9) using Matlab or Mathcad and using (5.7) to find the other parameter, the MLE method can produce estimates for  $k$  and  $\lambda$ . More results and comparisons of the different methods are present in section 5.1.4 below. Reference [22] contains much more detail on the theory and derivations involved for MLE

#### 5.1.3.4 Method of Moments

The method of moments is a widely used and convenient technique for estimating model parameters. The method is based on the observation that if two distributions have a certain number of moments in common, they will “look alike.” It can be assumed then that the set of moments of all orders uniquely determines the distribution[22]. In this method, the first  $m$  sample moments are equated to the corresponding population moments for the given measurement model. For many distributions, the population moments are a function of the parameter vector  $\theta$ , thus  $m$  parameter components are determined from  $m$  simultaneous equations [22]. This method can be applied to the gamma distribution.

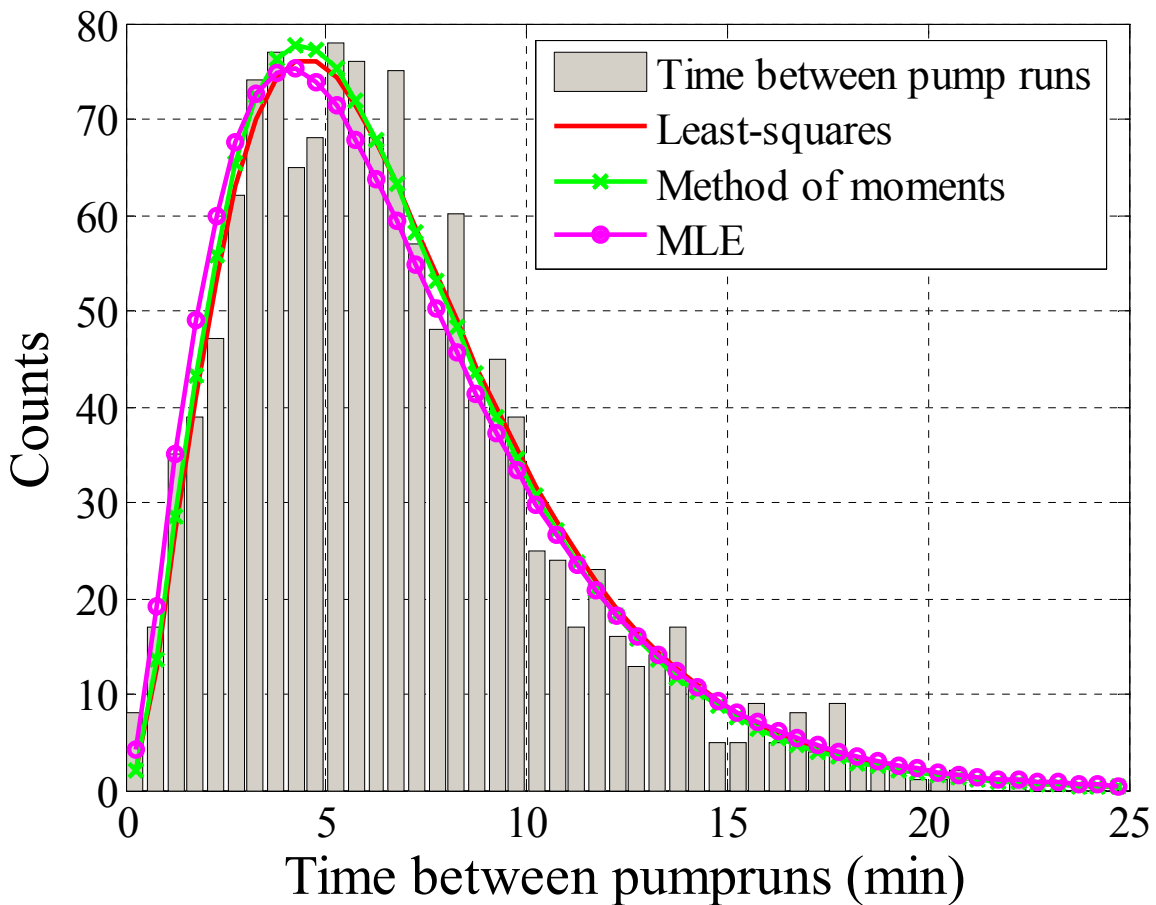
Using the first and second moments for the gamma distribution and solving the two moment equations simultaneously, the following relationships can be used to determine  $k$  and  $\lambda$  [22]:

$$k = \frac{\text{mean}^2}{\text{variance}} \quad (5.10)$$

$$\lambda = \frac{\text{mean}}{\text{variance}} \quad (5.11)$$

A comparison can be made for a curve fitted to Seneca data using the three different parameter estimating techniques. Using Matlab or Mathcad to perform a least squares fit, the fitted gamma PDF has the following parameters:  $k=2.887$  and  $\lambda=0.423$ . Using the method of moments, the parameter values were  $k=2.879$  and  $\lambda=0.431$  and MLE calculations showed  $k=2.566$  and  $\lambda=0.384$ . Figure 5-7 shows the resultant distributions

plotted over a week of Seneca data. Although each method clearly produces reasonable results, each has its own strengths and weaknesses. From an optimality standpoint, the maximum likelihood method is the most reliable estimation tool to apply in this case. As stated previously, least-squares can produce poor estimates when there are bins that contain few entries, and there clearly are in this case. Further, the method of moments can be shown to produce biased parameter estimates [23]. Regardless, the computational simplicity of the method of moments makes it an attractive approach. As shown in section 5.1.4, the results obtained using the MLE method and the method of moments are quite similar.



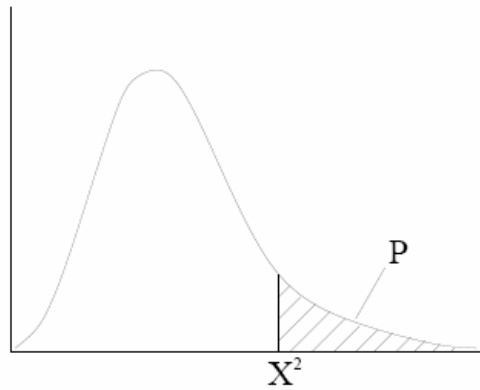
**Figure 5-7: Fitted Gamma distributions overlaid on Seneca data.**

#### 5.1.3.5 Goodness-of-fit Test

In order to make use of the parameter estimates obtained using the methods described above, a procedure must be developed to test how well the actual data fits the expected model. One simple analytic technique is to use the chi-squared ( $\chi^2$ ) goodness-of-fit test. Essentially, this procedure tests how well the binned data fit to the expected, normalized density function. In classical applications of the chi-squared test, two hypotheses are formed. One of these hypotheses, the “null hypothesis,” states that the expected distribution correctly describes the measured data. The other hypothesis, the “alternate hypothesis,” states that the data is not described by the current model. To determine which hypothesis is correct, we calculate the error between the expected distribution and the actual values. To quantify this error, we use the chi-square statistic, which is defined as

$$\chi^2 = \sum_1^M \frac{(O_i - E_i)^2}{E_i} \quad (5.12)$$

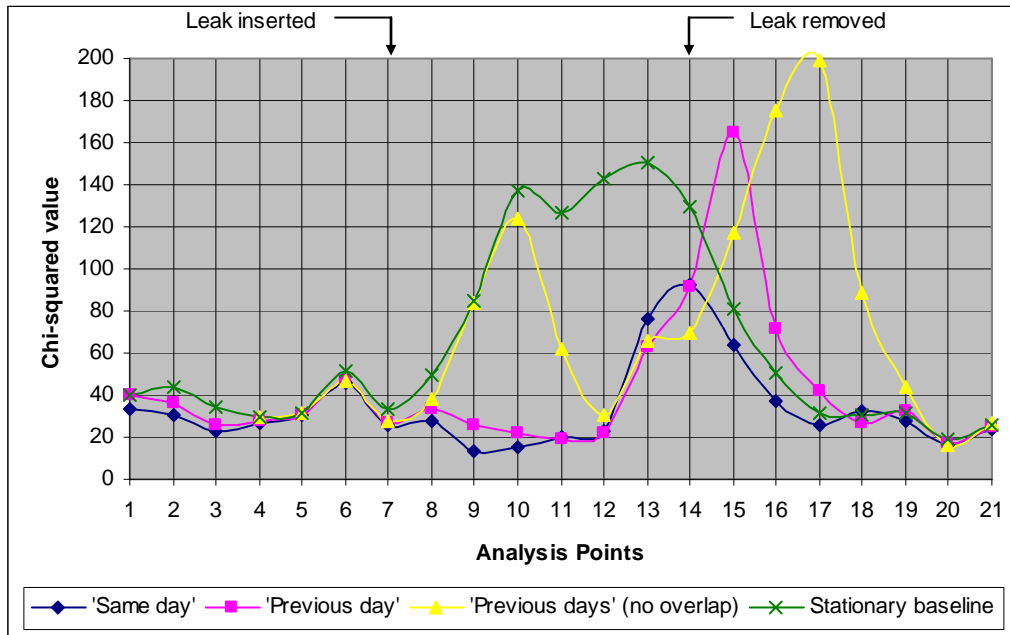
where  $M$  is the total number of bins and  $O$  and  $E$  denote the observed and measured values of the frequency distribution in  $i^{\text{th}}$  bin, respectively [20]. Note that the chi-square statistic is a random variable that is distributed according to the chi-square PDF. Essentially, this variable provides some indication of whether or not one can reasonably claim that the deviation between the actual data and the expected result is due to random chance. If the chi-square value becomes very large, it becomes increasingly less likely that the deviations are due to chance. Typically, one will choose a maximum allowed value for chi-square, and state that any chi-square values above that threshold correspond to data sets that do not fit the expected model. This procedure is illustrated graphically in Figure 5-8. A typical threshold is shown. Clearly, any values above that threshold should occur very infrequently if the model is true [20].



**Figure 5-8: Chi-squared distribution with typical threshold shown.**

When using the chi-squared test, it is necessary to determine an expected, or “baseline,” distribution. In this case, several possibilities were considered. One choice was to evaluate the gamma PDF using the parameters obtained during the estimation step. Additionally, we considered a “stationary” baseline using parameters obtained during a period when the system was known to have no leaks. We also compared data to what would be expected given the parameters obtained from data recorded several days prior to the current analysis period.

To demonstrate the use of a goodness-of-fit test a set of simulated data was generated for a 24 day period. During the first 10 days, no leak was present in the system. For the subsequent seven days, a leak was in place. Over the final seven days of the simulation, the leak was removed. At the end of each day, a diagnostic analysis was performed. For each analysis, least-squares was used to estimate  $k$  and  $\lambda$  for the histogram formed from the last three days of data; thus, each “analysis period” contains seventy-two hours of data. Figure 5-9 shows the results of several different chi-squared tests performed on the simulated data. Each of the chi-squared tests used a different baseline. The ‘same day’ baseline was a gamma PDF using the  $k$  and  $\lambda$  calculated for the current seventy-two hour period. The ‘previous day’ baseline used the  $k$  and  $\lambda$  value from the previous analysis period, and the ‘previous days (no overlap)’ baseline used the parameter values from three days prior (no overlap of data used to determine the two sets of parameters). Finally, the ‘stationary baseline’ used parameters from a period when the system was known to have no leaks.



**Figure 5-9: Various chi-squared analysis methods on three weeks of simulated data.**

As can be seen by the results, the chi-squared test against a stationary baseline is the best indicator of a leak. It has a distinct rise and fall corresponding to the times immediately following the introduction and removal of a leak in the system. The formation of a good stationary baseline is an issue, though, as there are regular fluctuations in system usage. Seneca data from August 2005 and from December 2005 demonstrates this point well. A week long sample taken from August (1297 total pump runs) has  $k=2.887$  and  $\lambda=0.423$  whereas a week long sample from December (1152 total pump runs) has  $k=2.94$  and  $\lambda=0.3942$ . There was no leak in the system during both periods. If the August baseline is used, the chi-squared value on the week in August is 59.585 while the chi-squared value on the week in December is 103.487. Reversing the process and using the December baseline, the chi-squared value on the week in August is 105.281 while the week in December is 64.181. The disparity between the resultant chi-squared values means that using one single baseline isn't going to be accurate and robust enough for all underway periods. Moreover, if system usage rises or falls significantly during any period, this method would clearly fail.

Based on the field observations made onboard Seneca and on knowledge of operations at sea, determining a robust baseline that would work in all situations and underway periods is not possible, thus the chi-squared method of determining the presence of a leak is not the best primary method of determining the health of the cycling

system. Other unreliability problems found with the chi-squared goodness-of-fit tests applied to Seneca data further showed that the test was not robust enough for implementation. During periods with leaks installed in the system, the resultant  $\chi^2$  values could be very low depending on how long the leak had been in the system. For instance, the results of some chi-squared tests performed on a system with a leak that had been present for five to seven days could not be distinguished from the results of tests performed on a system with no leaks.

#### *5.1.4 Parameter Trending*

With median filter and chi-squared analyses lacking enough robustness to be reliable in all situations, another method of determining system health is needed. Using both the estimated parameters values and the total number of runs, a fairly simple method exists that can detect a change in system status and determine if the change is due to a leak or to an increase in system usage.

In the case of increased system usage or in the presence of a leak, over a finite period of time the number of runs will increase and the mean time between pump runs will decrease. The distribution of times will narrow and shift to the left. When this occurs, the  $k$  and  $\lambda$  values associated with a fitted curve also change. Figure 5-10 and Figure 5-11 show the  $k$  and  $\lambda$  values for one week simulated leaks from 0 to 100 in-Hg/hour with  $\lambda=30$  flushes/hour and the same for 10 to 100 flushes/hour on a system with no leak.



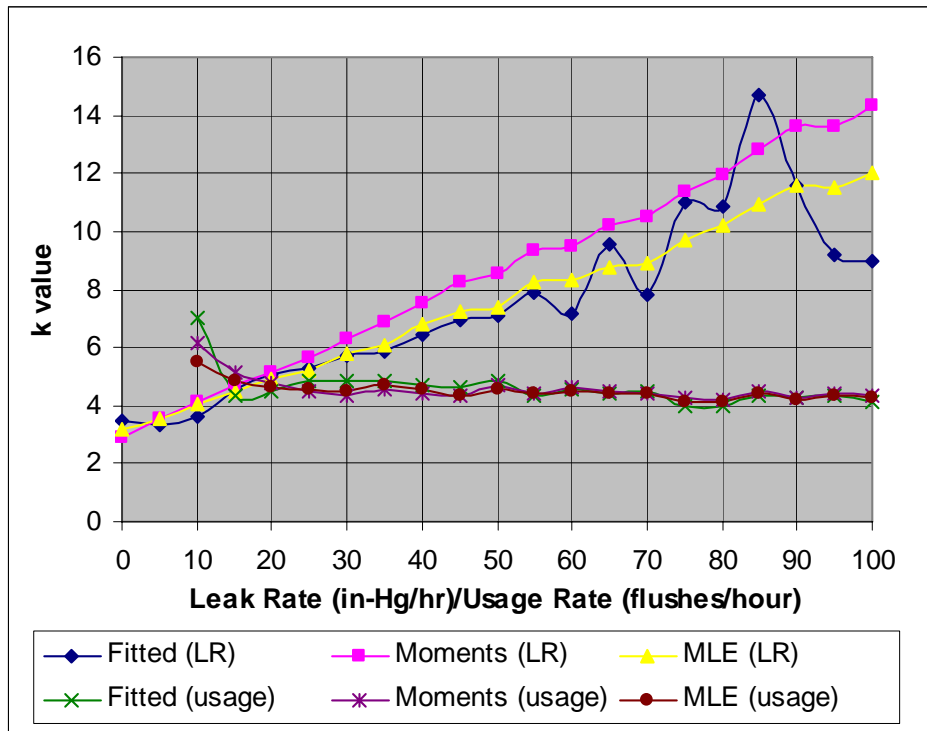


Figure 5-10: Gamma distribution fitted  $k$  values for varying leak rate (LR) and for varying usage rate (usage) using least squares, method of moments and maximum likelihood estimator methods.

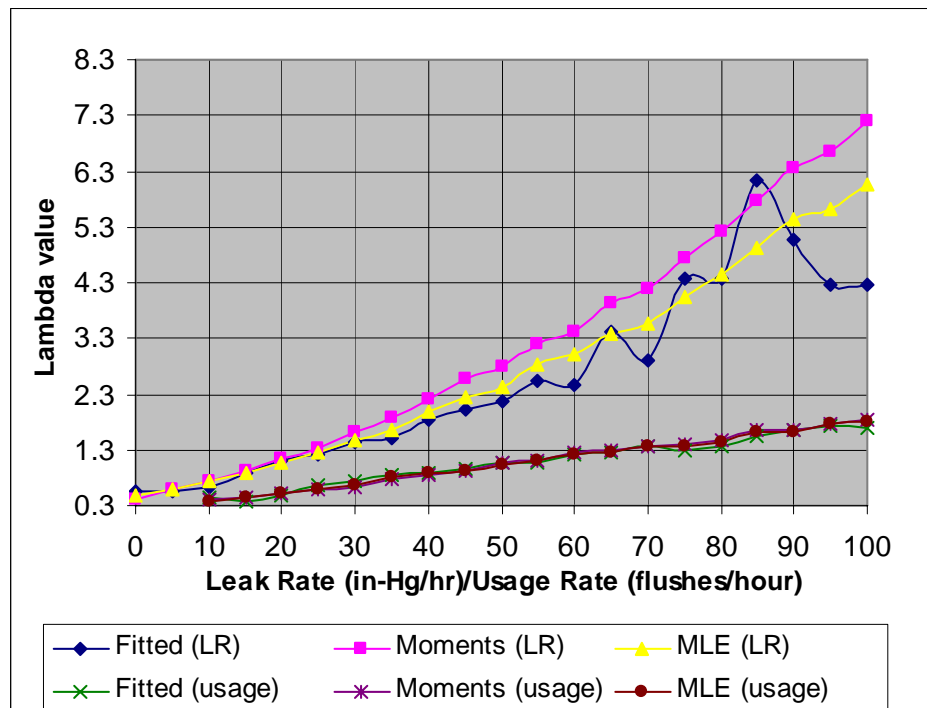


Figure 5-11: Gamma distribution fitted  $\lambda$  values for varying leak rate (LR) and for varying usage rate (usage) using least squares, method of moments and maximum likelihood estimator methods.

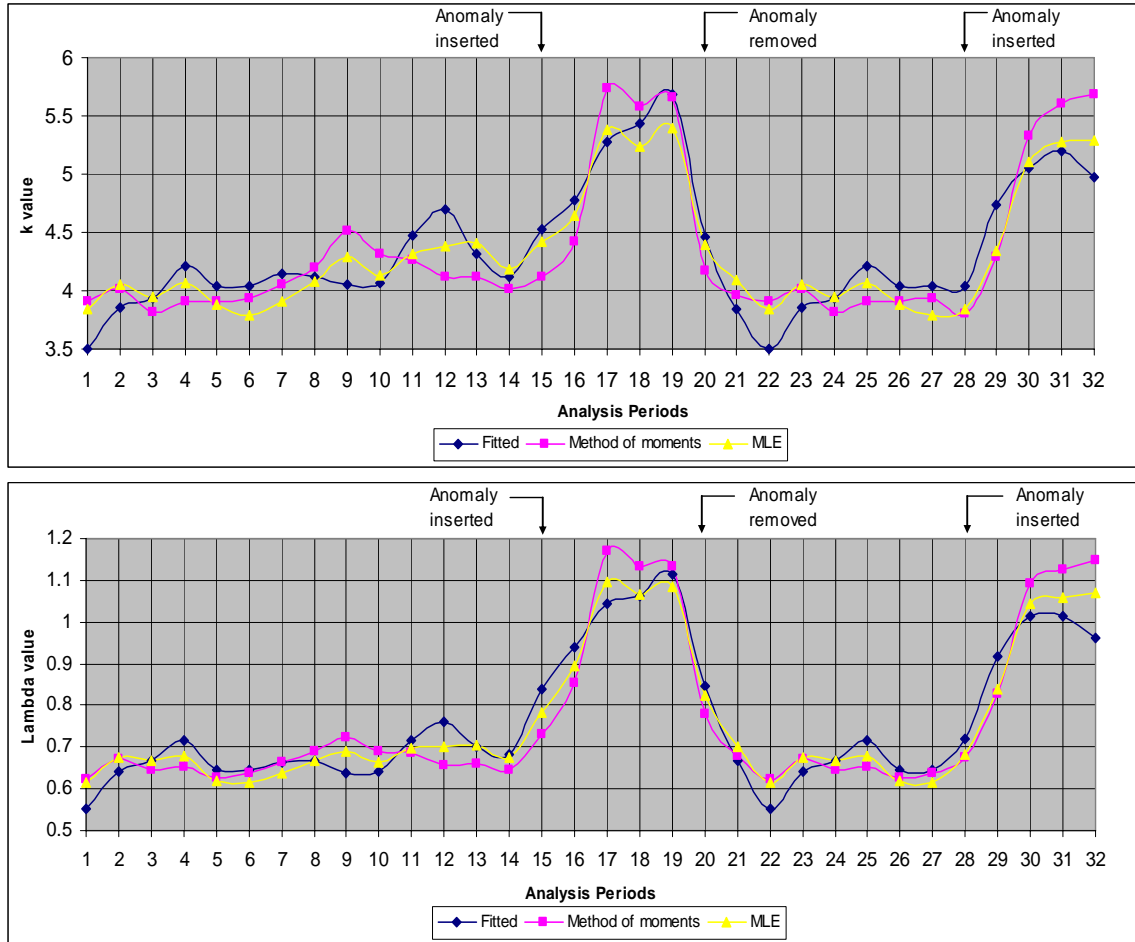
For both increased leak rates and for the increased usage rates, the total number of pump runs rose. The maximum number of runs achieved at 100 in-Hg/hr was 2774 and 2744 runs occurred when the usage rate was 100 flushes/hour. By analyzing the changes in the parameters, it can be seen that the  $k$  value remains relatively constant during increased usage periods but increases proportionally to the size of the leak. The  $\lambda$  value increases in both cases, but not as rapidly in the increased usage case. Using the fact that the  $k$  value is indicative of a leak in the system, parameter trending can be performed on simulated and real data.

#### 5.1.4.1 Simulated Data

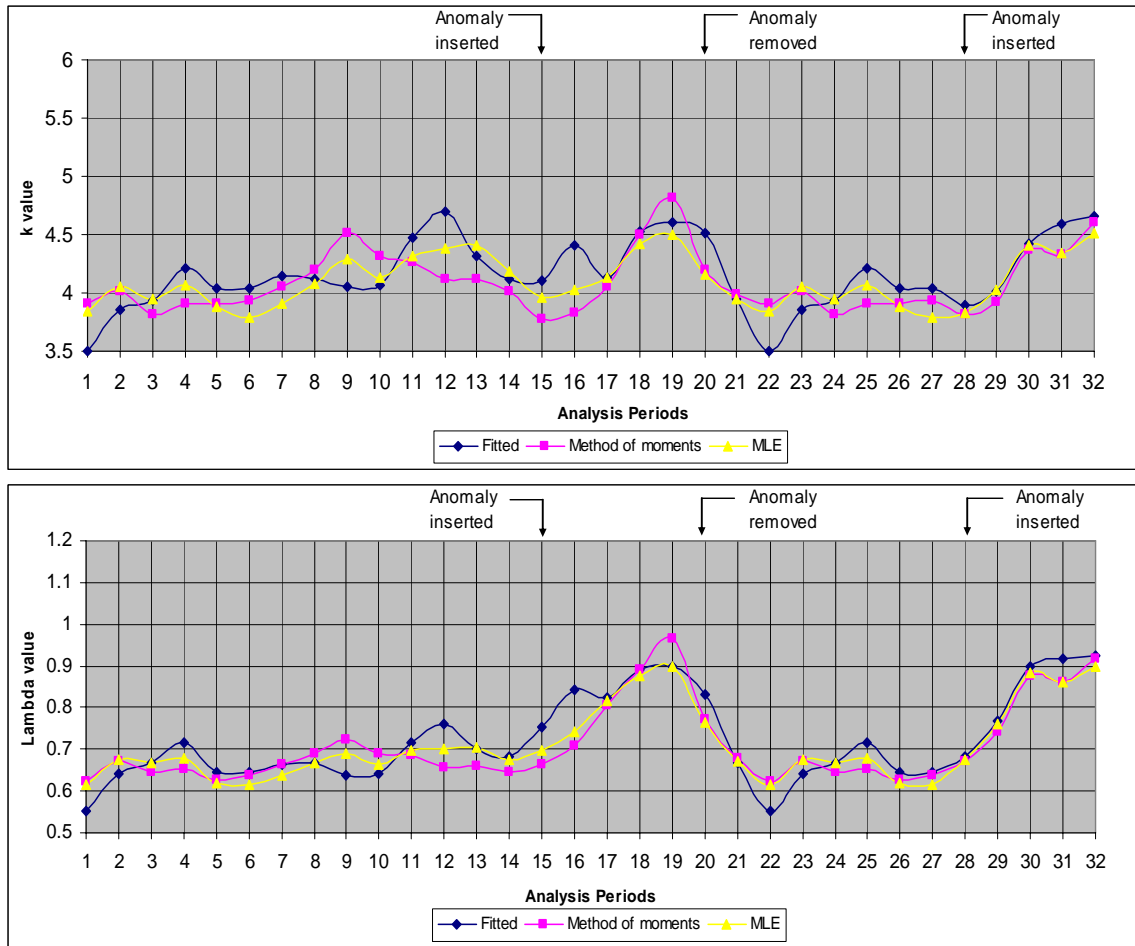
Simulated data was used first to test this method. Figure 5-12 and Figure 5-13 below show the  $k$  and  $\lambda$  values for the scenario described in Table 5-2. Each analysis period contains seventy-two hours of simulated data and each subsequent period contains forty-eight hours of overlap with the previous analysis period. The plots show the  $k$  and  $\lambda$  values calculated using the previously discussed methods.

**Table 5-2: Parameter trending data arrangement scheme for simulated data for inserted leak and increased usage rates.**

Analysis Period(s)	Inserted Leak(Figure 5-12)	Increased Usage (Figure 5-13)
1-14	No leak	$\lambda = 30$
15	48 hours: no leak 24 hours: 12 in-Hg/hour	48 hours: $\lambda = 30$ 24 hours: $\lambda = 40$
16	24 hours: no leak 48 hours: 12 in-Hg/hour	24 hours: $\lambda = 30$ 48 hours: $\lambda = 40$
17-19	12 in-Hg/hour	$\lambda = 40$
20	48 hours: 12 in-Hg/hour 24 hours: no leak	48 hours: $\lambda = 40$ 24 hours: $\lambda = 30$
21	24 hours: 12 in-Hg/hour 48 hours: no leak	24 hours: $\lambda = 40$ 48 hours: $\lambda = 30$
22-27	No leak	$\lambda = 30$
28	48 hours: no leak 24 hours: 12 in-Hg/hour	48 hours: $\lambda = 30$ 24 hours: $\lambda = 40$
29	24 hours: no leak 48 hours: 12 in-Hg/hour	24 hours: $\lambda = 30$ 48 hours: $\lambda = 40$
30-32	12 in-Hg/hour	$\lambda = 40$



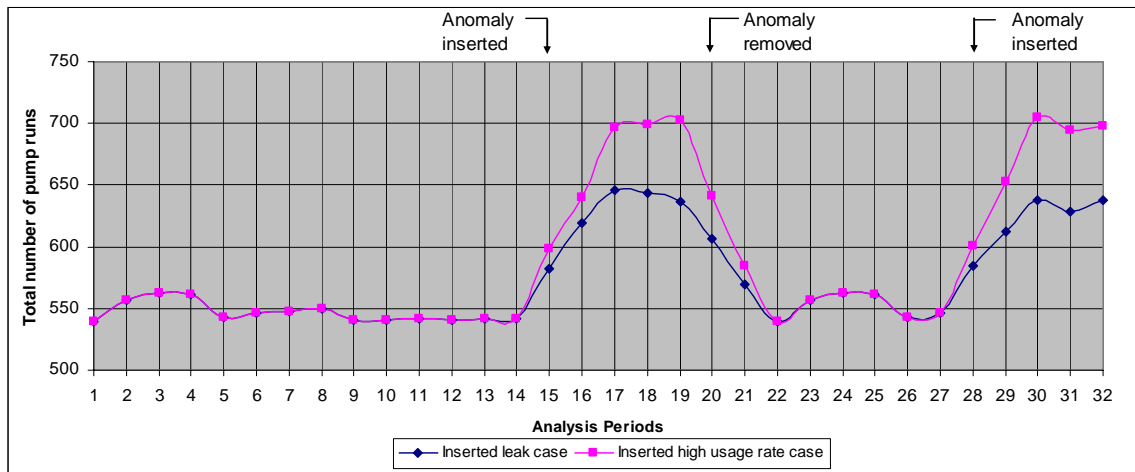
**Figure 5-12: Trended  $k$  and  $\lambda$  values for simulated data containing two periods with a 12 in-Hg/hour leak present.**



**Figure 5-13: Trended  $k$  and  $\lambda$  values for simulated data containing two periods of elevated usage (increase of 10 flushes/hour).**

As demonstrated by the plots, the  $k$  and  $\lambda$  values both increase in the presence of a leak. The  $k$  value rises to a value over 5.0 and the lambda value increases to over 1.0. On the other hand, when the increased usage occurs, the  $k$  value does not increase and the  $\lambda$  still rises. The increase in the  $\lambda$  value is less for the increased usage rate but is still evident.

A strong initial indicator for both of the above situations is the total number of runs in each seventy-two hour period. Figure 5-14 below shows the total number of runs for the above cases. Note that the increased usage rate results in a larger number of total runs.



**Figure 5-14: The total number of pump runs for the simulated cases.**

#### 5.1.4.2 *Seneca Data*

The parameter trending method was also tested on several sets of real data. Table 5-3 lists the conditions observed during a series of 30, 3-day analysis periods. The results of this test are shown in Figure 5-15 and Figure 5-16. Similar to the simulated data, both  $k$  and  $\lambda$  increased after the insertion of the leak. Additionally, note that the  $k$  value is not affected by the change in usage that occurs between point six and point eight.

**Table 5-3: Parameter trending data arrangement scheme for Seneca data for an inserted leak.**

Analysis Period(s)	Inserted Leak(Figure 5-16)
1-14	No leak
15	48 hours: no leak 24 hours: 12 in-Hg/hour
16	24 hours: no leak 48 hours: 12 in-Hg/hour
17-19	12 in-Hg/hour
20	48 hours: 12 in-Hg/hour 24 hours: no leak
21	24 hours: 12 in-Hg/hour 48 hours: no leak
22-27	No leak
28	48 hours: no leak 24 hours: 12 in-Hg/hour
29	24 hours: no leak 48 hours: 12 in-Hg/hour
30	12 in-Hg/hour

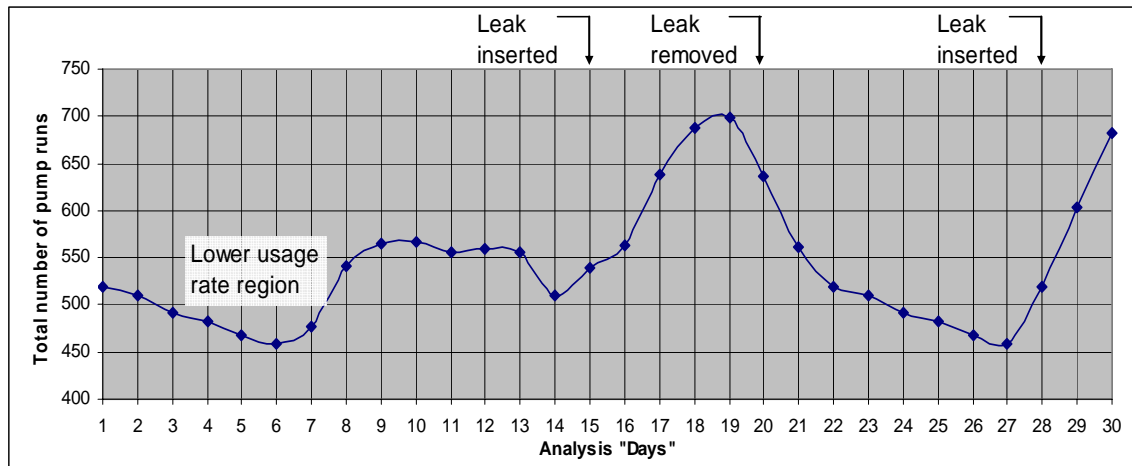


Figure 5-15: Total number of pump runs for Seneca data parameter trending periods.

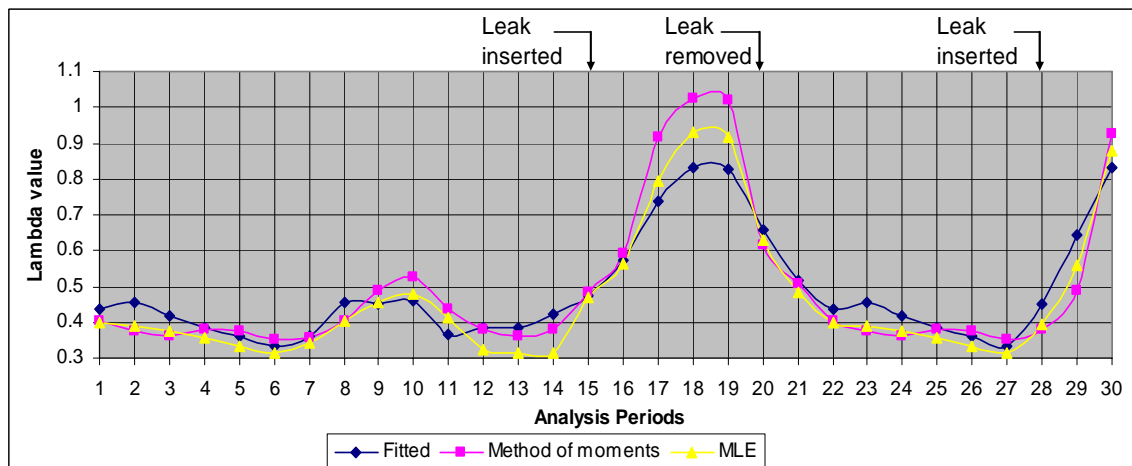
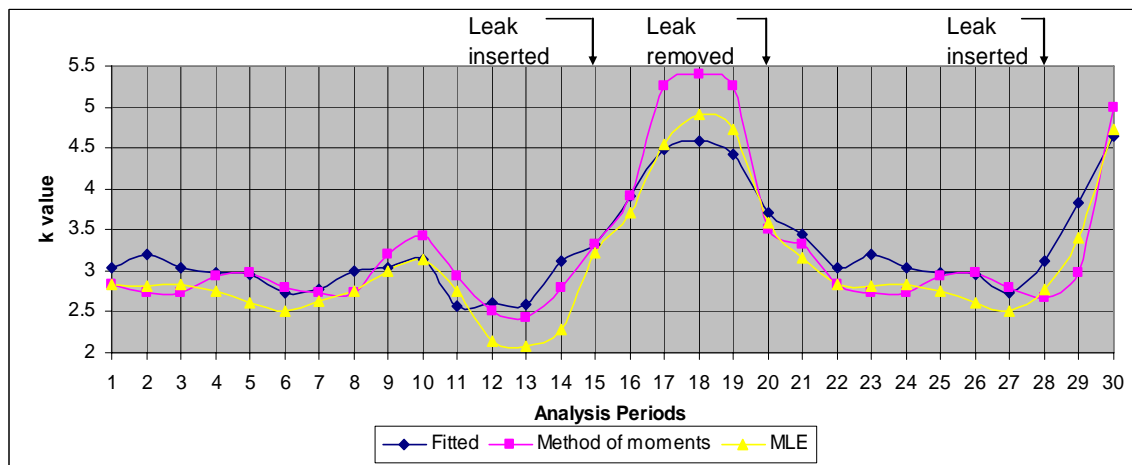
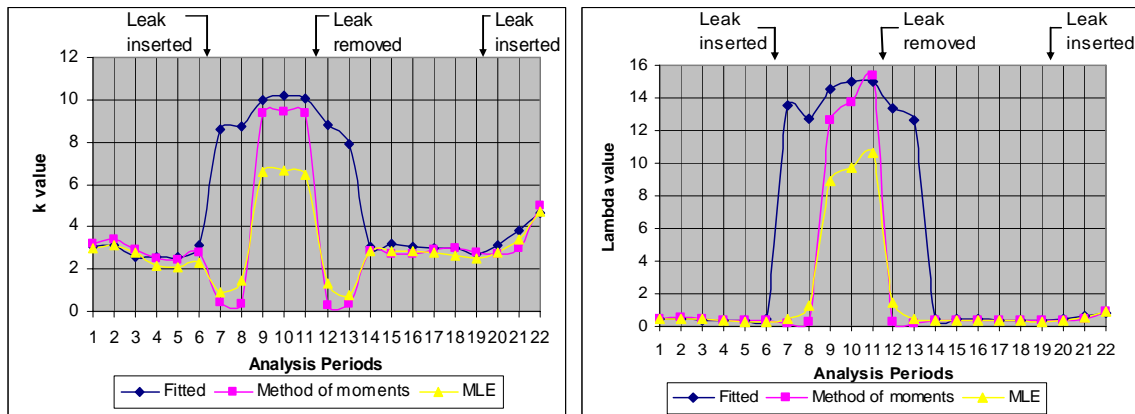


Figure 5-16: Seneca  $k$  and  $\lambda$  values results for inserted leak cases.

The method was also tested during a period in which a massive leak occurred due to a check valve failure. The massive leak starts at analysis point seven and is removed at

point eleven. A smaller, 12 in-Hg/hour leak was inserted at the end of analysis period nineteen. As can be see in Figure 5-17, the fitted  $k$  and  $\lambda$  values rapidly and greatly rise after the leak is inserted. The method of moments and MLE break down immediately after the introduction and removal where the calculated  $k$  value drops to nearly zero. It remains near zero while the analysis periods contain a mixture of data from both the major leak and from the no leak condition. Once the analysis period contains data from only the no leak or the major leak condition, the method of moments calculations then result in  $k$  and  $\lambda$  values that are similar to the fitted values.

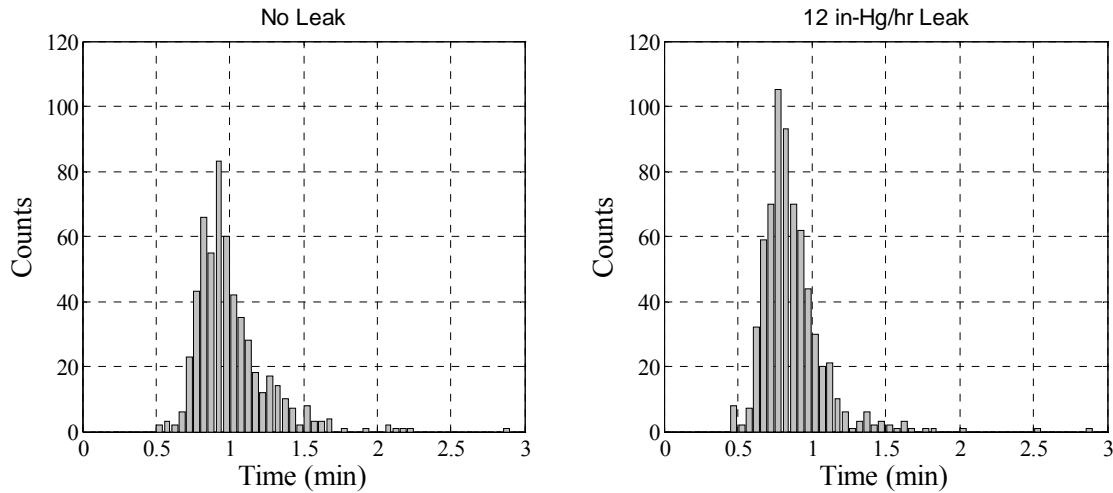


**Figure 5-17: Trended  $k$  and  $\lambda$  values for Seneca data with check valve failure.**

### 5.1.5 Load Time Analysis

The previous analyses have all looked at the time between pump runs. The length of time that the pumps run on each cycle can also be examined. Figure 5-18 below shows the loaded run times for three days of underway Seneca data with no leak and three days with a 12 in-Hg/hour leak inserted. The mean run time with no leak in the system was 1.0048 minutes while the mean loaded time with the leak inserted was 0.8664 minutes. This seems counterintuitive but it is expected. For the no leak condition, the low pressure setpoint and subsequent pump start is more often caused by a flush event. When the flush occurs, the system pressure drops below the low pressure setpoint, thus the pump must operate long enough to not only raise the pressure from the low pressure setpoint to the high pressure setpoint, but also long enough to initially raise the pressure from the final pressure after the flush up to the low pressure setpoint. With the leak installed, more often the low pressure setpoint is reached merely by the reduction of pressure from

the leak, thus the pump, once energized, only has to raise the pressure from the low pressure setpoint to the high pressure setpoint.



**Figure 5-18: Loaded run times for Seneca vacuum pumps with no leak condition (left) and 12 in-Hg/hour leak condition (right).**

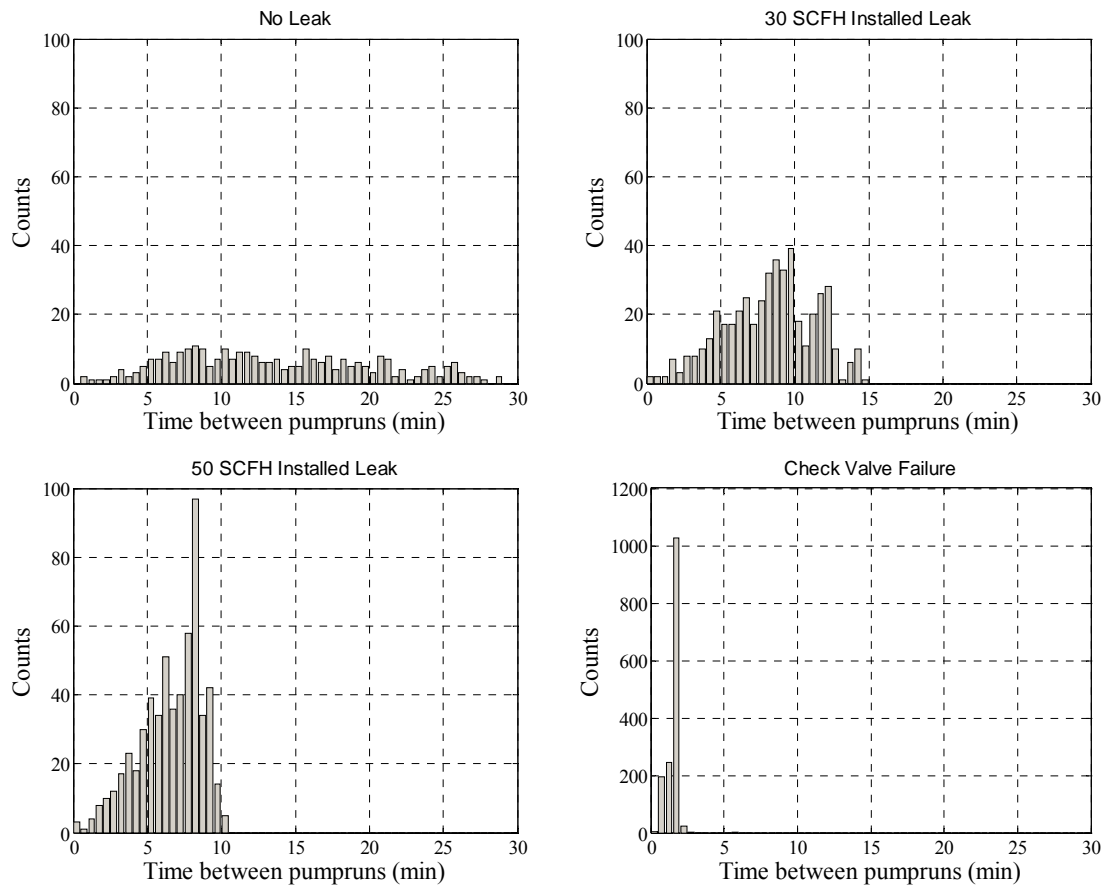
### 5.1.6 Inport Data Considerations

The previous discussions have been using  $\lambda$  values of thirty or more, which is representative of underway usage rates. Since the crew has no option except to use the facilities onboard the ship while underway, it is expected that at sea usage rates would be greater than inport rates. When the ship is in home port, approximately eighty-five percent of the crew goes home at night and, although onboard the ship for a typical work day, the crew tends to not use the restrooms as often as they would at sea.

A detection method used inport has to discriminate between an elevated usage rate and a leak, just as at sea; however an elevated usage rate is not likely to occur for all twenty-four hours during the day. Due to the fact that usage is both low and rather sporadic, the Poisson model has not been found to be an accurate model for inport behavior. As a result, parameter estimation techniques cannot be applied in this case.

Since the Seneca is more accessible for experimentation during inport periods, numerous leak rates and experiments have been performed. Using the same seventy-two hour data grouping period, a typical histogram of times between pump runs is shown in Figure 5-19 along with histograms from periods with installed leaks in the system and a period during check valve failure.





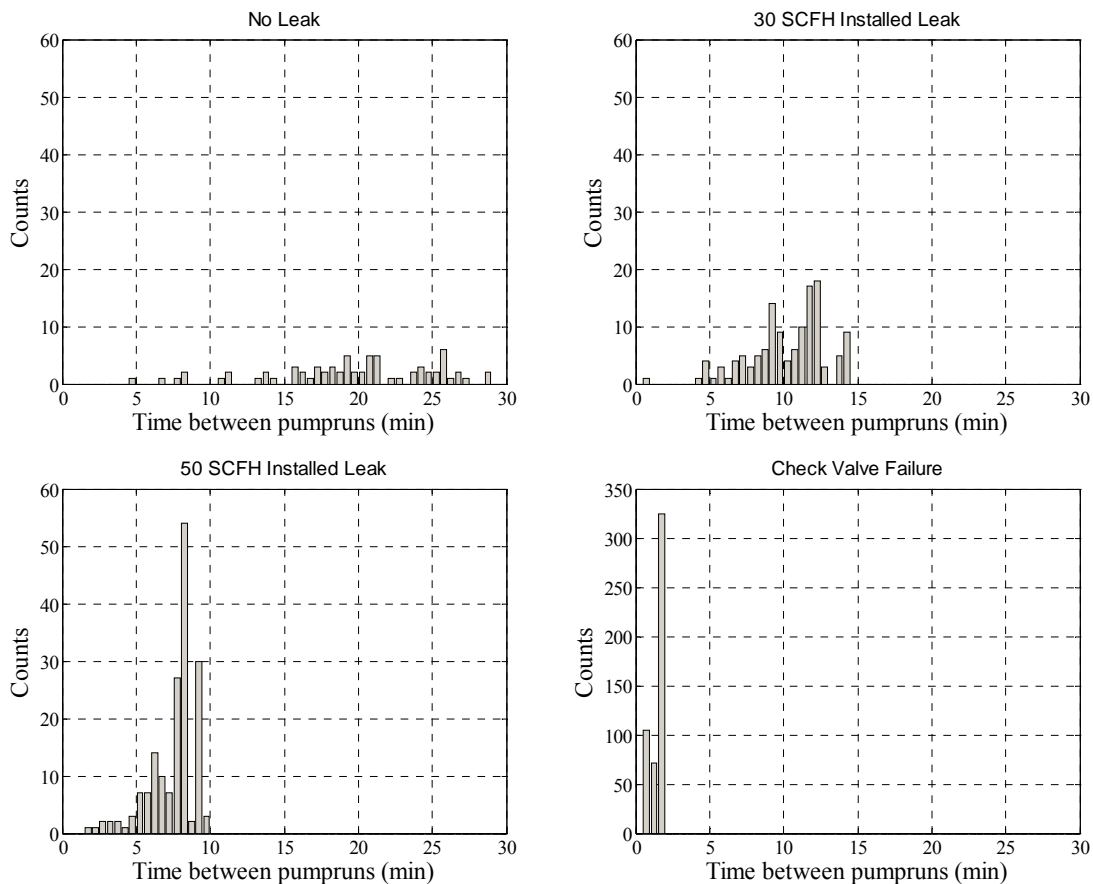
**Figure 5-19: Inport Seneca data for no leak, 30 SCFH leak, 50 SCFH leak, and check valve failure.**

In order to install a leak, an air flow meter was installed in the system and was adjusted as necessary to get the desired leak rate. Two of the plots of Figure 5-19 show the histograms for seventy-two hour, inport periods with 30 SCFH and 50 SCFH installed leaks. Table 5-4 shows the number of runs and mean times between pump runs for the four different conditions: no leak, 30 SCFH, 50 SCFH, and during check valve failure. As evidenced by the times and numbers, a leak is easy to detect.

**Table 5-4: Total number of runs and mean times between pump runs for Seneca inport periods.**

<u>Leak Rate</u>	<u>Number of pump runs</u>	<u>Mean time between runs</u>
None	295	13.652 minutes
30 SCFH	467	8.283 minutes
50 SCFH	575	6.59 minutes
Check Valve Failure	1498	1.51 minutes

For an increased number of pump runs, the best scheme to determine if a leak exists is to examine the nighttime data (Figure 5-20). Because of the unique situation where the usage rate during the nighttime is extremely low, any leak will manifest itself in a more narrow distribution, and the bin time associated with the right edge of the distribution will correspond to the time it takes for the system pressure to drop from the high pressure setpoint to the low pressure setpoint without any flushes occurring. For instance, if the rightmost bin is located at ten minutes, then the approximate leak rate is 24 in-Hg/hour because the system pressure decreased 4 in-Hg in  $1/6^{\text{th}}$  of an hour.



**Figure 5-20: Inport nighttime Seneca data for no leak, 30 SCFH leak, 50 SCFH leak, and check valve failure.**

Based on the longest time reached during the nighttime period, a 30 SCFH leak corresponds to a 17 in-Hg/hour leak rate, a 50 SCFH corresponds to a 25 in-Hg/hour leak rate, and the check valve failure condition corresponds to greater than 130 in-Hg/hour leak rate.

Anecdotally, one inport measurement period taken after the ship was inport for a couple of months revealed a very large leak. The crew did not have any idea that a leak

was in the system since no alarms or other warnings existed to alert them to the problem. The histograms of times taken over a seventy-two hour periods are included in Figure 5-19 and Figure 5-20. The leak rate associated with the histogram was approximately 135 in-Hg/hour.

The author alerted the crew to the leak and along with a few members of the crew discovered the two check valves at the suction of the vacuum pumps to be faulty. The check valves are meant to shut after the vacuum pump de-energizes and maintain the vacuum in the system. Disassembly of the valves revealed pitted faces and loose components. The valves were beyond repair and had to be replaced. After replacement, the histogram showed a great improvement. The figures below show the condition of the valve and internals as it was disassembled.



**Figure 5-21: Photos of failed check valves: as opened (upper left), pitted valve face (upper right), rubber valve face with uneven wear marks (lower left), and pitted face of second valve (lower right).**

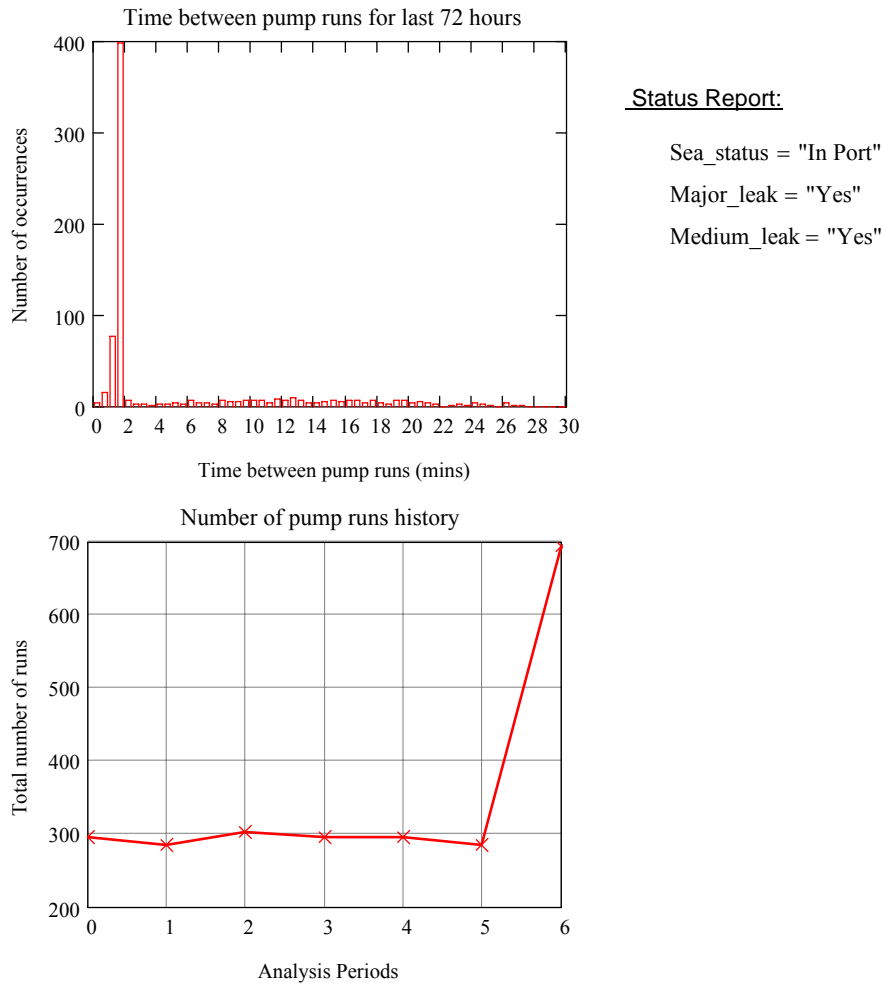
## 5.2 Diagnostic Method and Status Reports

Using the methods above, two diagnostic schemes were developed. One scheme is used for inport conditions and one for underway conditions. In both cases, there is a “one-day” indicator designed to detect a massive leak and a “three-day” indicator to detect small to medium leaks where the “one-day” indicator utilizes the previous twenty-four hours of data and the “three-day,” the previous seventy-two.

Detecting the transition between inport and underway is not an easy task to perform automatically. If there are no leaks in the system, the number of runs per day, or a change in the number of runs per day, can be a good indication the ships status. However, if a leak arises during an inport period, the increased number of runs from the leak can falsely lead the NILM to assume that the ship has gotten underway. For the initial software development, a report of the current system status will require an “At sea or inport?” input from the user. The “inport” calculations and the “at sea” calculations will run simultaneously at all times with two status reports will always be available. Selection of the ships status by the user will produce the report applicable to the ships condition.

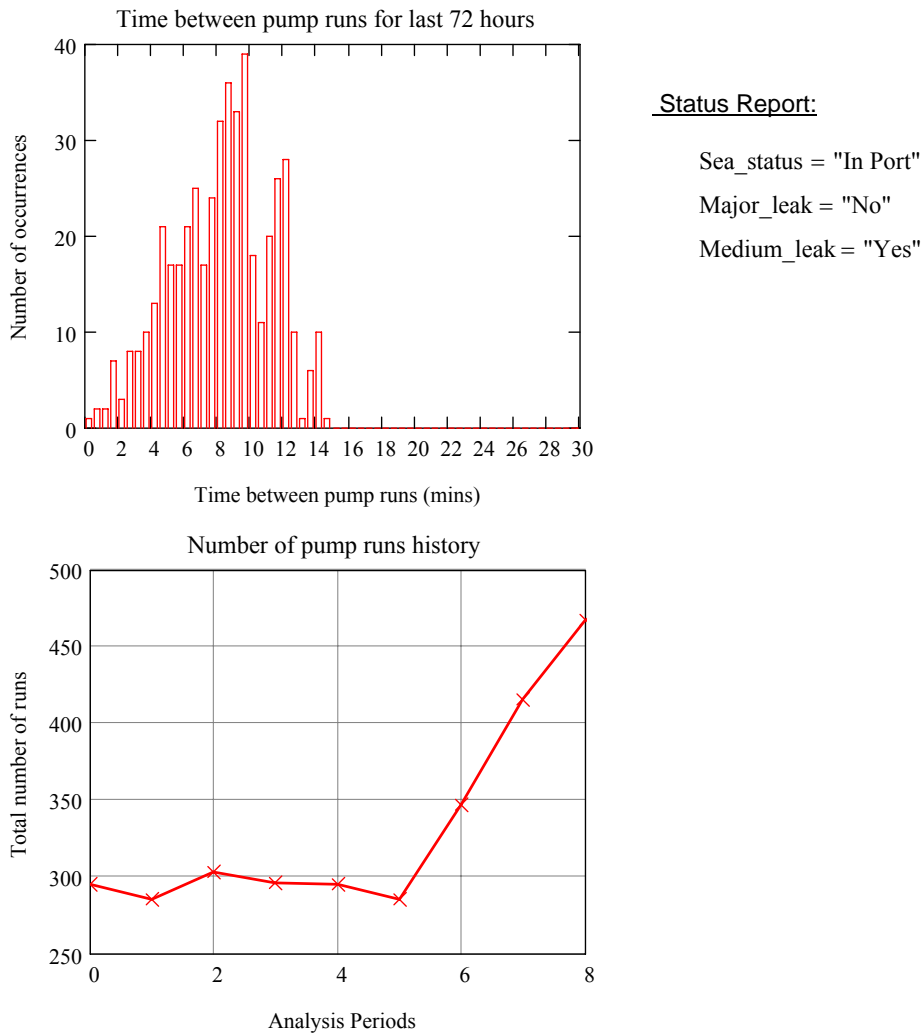
### 5.2.1 *Inport Diagnostic Method*

The “one day” indicator for the inport periods will monitor both the total number of runs per twenty-four hour period and the results of a median filter test. If the total number of runs were to increase by more than 200 and if the maximum difference between the binned data and the median filtered data was over 150, then the status report would indicate the presence of a major leak. Figure 5-22 shows a representative status report that was generated within twenty-four hours after the introduction of a leak while the Seneca was inport.



**Figure 5-22: Inport status report twenty-four hours after check valve failure.**

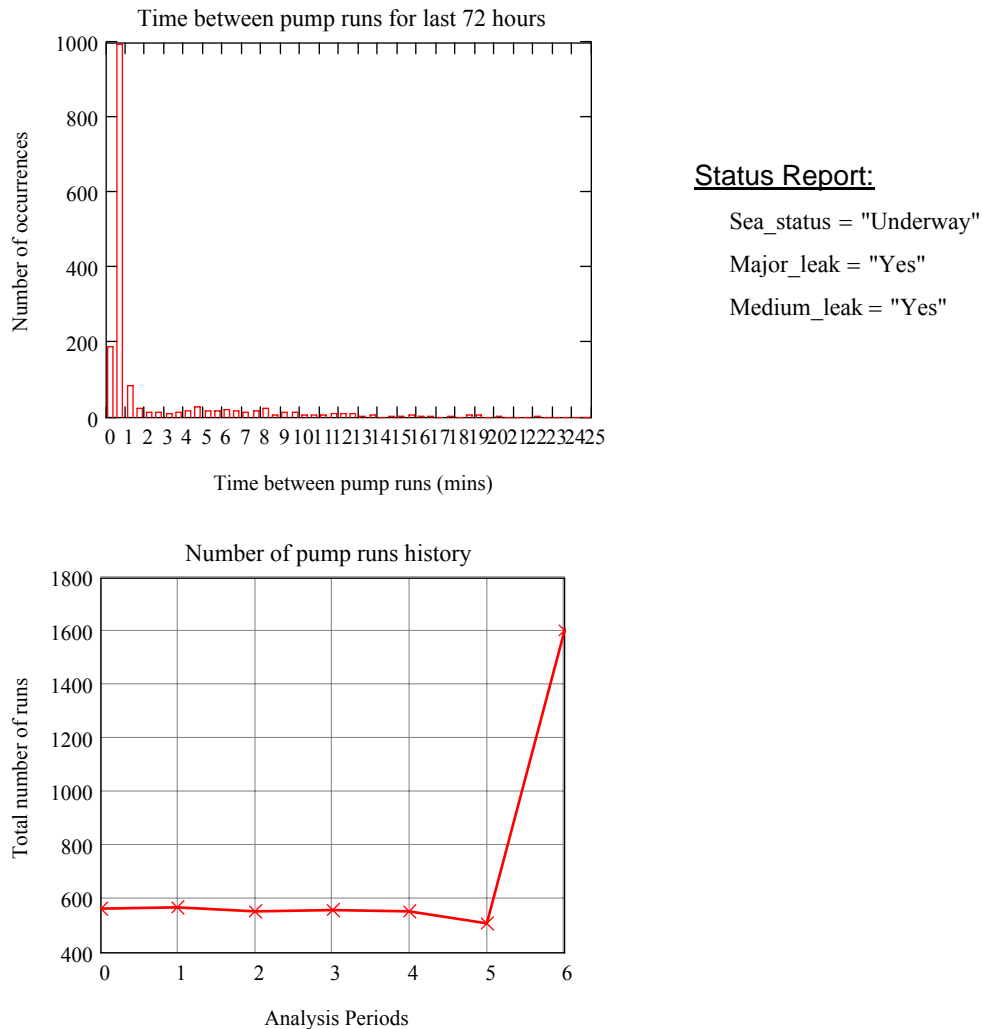
The “three-day” indicator for the inport periods will consist of three checks. First, the median of the times associated with the five bins that have the longest time between runs would be monitored. If that median time is less than twenty minutes, a leak is likely. Second, an edge detection scheme would also evaluate the histogram data to detect any sharp edges in the histogram. Lastly, the total number of runs would be monitored and compared to the average over the previous three analysis periods. An increase over the average of seventy-five runs is a strong indicator of a system change. If two out of the three “three-day” tests indicate the presence of a leak, the user will be alerted on the status report. Figure 5-23 shows a representative status report produced within seventy-two hours after introducing a 17 in-Hg/hour leak into the system.



**Figure 5-23: Inport status report seventy-two hours after insertion of 17 in-Hg/hour leak.**

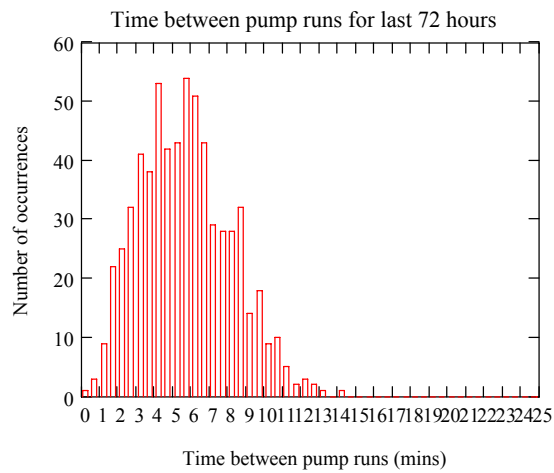
### 5.2.2 Underway Diagnostic Method

Similar to the inport diagnostic method, the number of pump runs per day is monitored for increases over 200 and a median filter test is used to detect any abnormal discontinuities. Together, these “one day” tests can give an indication of a major leak such as that from a check valve failure. Using Seneca underway data, the status report shown in Figure 5-24 was generated within twenty-four hours after the valve failure.



**Figure 5-24: Underway status report twenty-four hours after check valve failure.**

For the “three-day” tests, the k value is trended and monitored for levels over 4.25, a level which has previously shown to be an indication of a leak. If the k value increases over that threshold, then the number of runs is compared to an average of the last three totals, and if the total number of pump runs has increased over seventy-five runs, then the medium leak status is changed to “Yes.” An edge detection routine is performed on the underway data as well to find large discontinuities. Underway Seneca data was used to demonstrate the test and Figure 5-25 shows the status report generated within seventy-two hours after the introduction of a 12 in-Hg/hour leak.

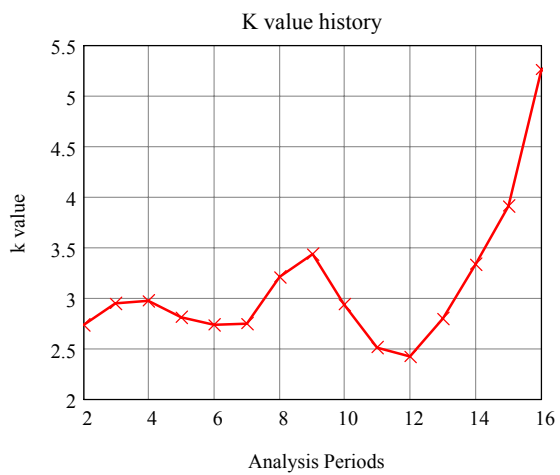
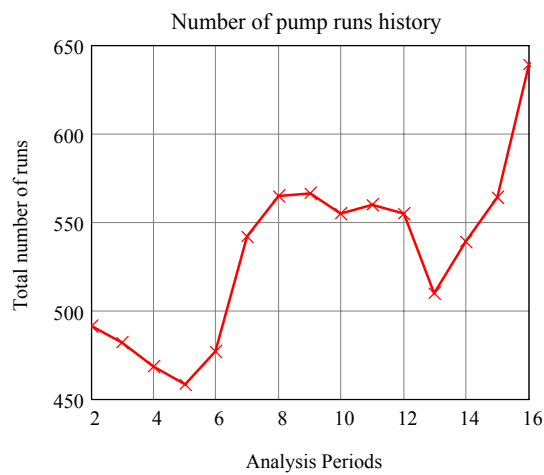


**Status Report:**

Sea\_status = "Underway"

Major\_leak = "No"

Medium\_leak = "Yes"



**Figure 5-25: Underway status report seventy-two hours after introduction of 12 in-Hg/hour leak.**



## **6 Cost Analysis for Monitoring of Seneca Sewage System**

### **6.1 Motivation**

Since space and weight are key considerations in any ship design, any equipment placed onboard has to have a purpose and a function necessary for the operation of the ship systems. As discussed previously, a NILM installed onboard a ship is intended to replace sensors within an engine room and without. Monitoring multiple pieces of equipment with one sensor vice multiple sensors is absolutely necessary to reduce the number of sensors on ship systems. The NILM is ideally suited to do just that with the capability to perform diagnostics not only on just the electrical equipment but also on the mechanical aspects of the systems.

Of equal importance to space and weight is cost. There are many facets of cost estimation and cost benefit analysis. The analysis presented below is not intended to be a detailed analysis, but a basic demonstration of the potential value that a NILM can add to a system. The specific cases used as examples are anecdotal cases onboard USCGC Seneca. Some precise costs are not publicly available, are very difficult to obtain, or are constantly changing, so the numbers used are best estimate speculations based on the information available.

Although the NILM monitoring the sewage system onboard Seneca currently only monitors four pumps (two vacuum and two discharge), the potential benefits of using a NILM to monitor multiple systems is also explored.

Medium leaks were inserted during testing and large leaks occurred during check valve failures. One check valve occurred at sea during October 2003 and November 2003 and the second occurred inport during February 2006. During both occurrences, the engineering crew did not immediately know of any problems associated with the sewage system because of the lack of monitoring on the system.

## 6.2 Power Calculations

Figure 2-8 shows a typical power trace for a vacuum pump. As can be seen by the scale of the plot, the approximate power level of the vacuum pump is slightly less than 2 kW. By examining the power plots during various underway and inport times, an average power level can be calculated. An averaged power level takes into account the large initial power spike and the slight undershoot as the power steadies out during a normal operation. The two vacuum pumps actually have slightly different power signatures, with one pump power average approximately 1.8 kW and the other near 1.7 kW. The pumps alternate on subsequent pump runs, so an overall average power level takes into account the variation between pumps.

The power was examined for four different situations onboard the Seneca and the calculations are included in Table 6-1. The first column shows a no leak condition inport. A smooth 300 minute sample uninterrupted by a discharge pump run or any other anomaly was examined. Over the 300 minutes, the number of pump runs, the total time the pumps were de-energized, and the total energy used (in kW-min) were obtained from the data. The time energized was the difference of 300 minutes and the time de-energized. Taking the total energy used divided by the total time energized gave the average vacuum pump power. The same calculations and examinations were done for three other conditions: a no leak condition at sea, a check valve failure inport, and a check valve failure at sea.

**Table 6-1: Average pump power level calculations.**

	<u>No Leak/Inport</u> <u>(11/11/2005)</u>	<u>Check Valve Failure</u> <u>Inport (02/14/2006)</u>	<u>No Leak/At Sea</u> <u>(08/18/2005)</u>	<u>Check Valve Failure</u> <u>At Sea (10/31/2003)</u>
Sample time: (min)	300	300	300	300
Total number of pump runs:	21	103	44	336
Total time pumps de-energized: (min)	280.895	167.784	259.678	182.448
Total time pumps energized: (min)	19.105	132.216	40.322	117.552
Total energy used: (kW-min)	32.238	241.118	71.420	219.217
Average pump power level (kW)	1.687	1.824	1.771	1.865

The average power level is 1.79 kW for these conditions. To demonstrate the amount of excess energy expended during a check valve failure condition, a simple

calculation can be done to show what happens over thirty days of undetected valve failure. By examining the inport data over extended periods of time (at least three days for each), an average amount of “up time” or time that a pump was energized (in min/day) was calculated for each of the four conditions above. By multiplying the “up time” per day by thirty days and by the average power level for each pump, a total energy expenditure was calculated. Table 6-2 shows the results. If inport, the kW-hrs are supplied by shore power, so the approximate cost of one kW-hr is used [17]. If at sea, the kW-hrs are supplied via the diesel generators, so the approximate electrical plant efficiency, the average specific fuel consumption (sfc) of the diesel and the cost of a gallon of fuel [18] are used to determine the cost of the extra energy expended. The two resulting numbers are examples of the excess costs over normal usage costs that could result from a month of a major undetected system leak. Ignoring the wear and tear on the pumps during the excessive operations, the excess cost of a month-long undetected leak can approach \$100.

**Table 6-2: Excess energy costs associated with a 30-day undetected check valve failure condition.**

<u>Result</u>	<u>No Leak Inport</u>	<u>Check Valve Failure Inport</u>	<u>No Leak At Sea</u>	<u>Check Valve Failure At Sea</u>
Avg. time pumps running (min/day)	101.330	647.150	206.250	689.610
Energy expended in 30 days (kW-hr)	90.690	579.199	184.594	617.201
Excess energy over no leak (kW-hr)		488.509		432.607
Cost per kW-hr (\$/kW-hr)		0.143		0.211
Cost of excess energy expended (\$)		69.86		91.29

Anecdotally, it is not unreasonable for such a large leak to go unnoticed for a month. In the most recent check valve failure case, the leak went undetected by the crew for four weeks before analysis of NILM data showed that a major leak was in the system.

### 6.3 Cost-Benefit Analysis

In order to assess the value of a NILM installed in the system such as the sewage system, a long term assessment was done. By examining the costs associated with two different sewage system “lifetimes,” the benefit of a NILM can be shown. Seneca was commissioned in 1987 and is expected to be decommissioned in 2025, so a fictional ship built in 2006 would be decommissioned thirty-eight years later in 2044.

One of the “lifetime” comparisons is for a sewage system with a NILM installed on it. The following assumptions were made about this system:

- The one NILM installed monitors only the sewage system and no other systems.
- The PC based computer has to be refreshed every five years, but the NILM hardware lasts fifteen years before it needs to be replaced.
- Every sewage system leak is detected by the NILM and the crew rapidly finds and fixes the problem.
- The vacuum pump and motor last the lifetime of the ship without need of replacement.
- The cost of shore power, fuel and replacement computer parts follow average inflation rates.
- The initial purchase and installation of the sewage system is not included in the lifetime costs.

The other “lifetime” comparison is for a sewage system without a NILM installed. The assumptions associated with this system are:

- A major leak occurs once every two years because of check valve failures (based on recent Seneca occurrences).
- A medium sized leak occurs every other year (opposite to the check valve failure years).
- Each of the leaks goes unnoticed for thirty days and is fixed upon discovery.
- The vacuum pumps and motors must be replaced after thirty years because of wear from the excessive starts associated with the leaks.
- The cost of shore power, fuel and replacement parts follow average inflation rates.
- The initial purchase and installation of the sewage system are not included in the lifetime costs.

Table 6-3 lists the assumptions used in the calculations for the two different lifetimes. The amount of time spent inport and at sea is based on the U.S. Coast Guard LANTAREA target operational schedule. The average amount of time the pumps run per day for the various situations is based on Seneca data. A month of non-detection was assumed for each of the leaks and the average pump power level calculated above was assumed for all the scenarios. Efficiency of the electric plant and the specific fuel consumption of the diesel generators are typical values and not based on Seneca equipment. The energy prices are based on the current rates for the Seneca and the

computer and NILM costs are all approximate costs. Lastly, the cost of a new vacuum pump was obtained from the U.S. Coast Guard repair parts supply system.

**Table 6-3: Inputs used in cost-benefit model to determine lifetime costs of shipboard sewage system.**

0.5	Percentage of time spent inport
0.5	Percentage of time spent at sea
101.33	Average time pumps run per day inport with no leak (min/day)
206.25	Average time pumps run per day at sea with no leak (min/day)
145.9	Average time pumps run per day inport with medium leak (min/day)
226.7	Average time pumps run per day at sea with medium leak (min/day)
647.15	Average time pumps run per day inport with check valve failure (min/day)
689.61	Average time pumps run per day at sea with check valve failure (min/day)
0.08	Fraction of year that a medium leak occurs
0.08	Fraction of year that a check valve failure occurs
1.76	Average power level of pumps (kW)
0.95	Efficiency of the electrical plant
0.29	Specific fuel consumption of diesel generator (kg/kW-hr)
1	Number of systems monitored by the NILM
2.25	Cost of 1 gallon of DFM (2006\$)
0.146	Cost of 1 kW-hr inport (2006\$)
1000	Cost of one new NILM unit (2006\$)
300	Cost of PC upgrade/rebuild (2006\$)
700	Cost of NILM hardware upgrade/rebuild(2006\$)
11318	Cost of one new vacuum pump (2006\$)
0.03	Average inflation rate
0.05	Average discount rate

The most expensive component of a NILM system as it is currently assembled is the PCI data acquisition card. Current costs for the card are on the order of \$600 [19]. The card is necessary but the NILM only uses a fraction of the capability of the card. A new data acquisition card has been specifically designed for NILM applications [14]. Although the exact cost of the new card is not available, estimates place it at less than \$100 per card. Assuming that the new card is used instead of the current PCI card, the cost of a new NILM unit and the NILM upgrade can be reduced by \$500 each.

Another potential cost saving situation is when the NILM monitors more than one system. Assuming that the same NILM monitors two systems, the cost of the install is spread to both systems, thus the cost associated specifically with the sewage system is cut in half.

These costs savings measures each result in new “lifetime” costs. Table 6-4 below shows a comparison of the results.

**Table 6-4: Discounted lifetime costs for shipboard sewage system with and without NILM installed.**

<u>NILM configuration</u>	<u>with NILM</u>	<u>without NILM</u>
One system monitored, current data acq. card	\$11,735	\$22,218
One system monitored, new data acq. card	\$10,580	\$22,218
Two systems monitored, current data acq. card	\$10,048	\$22,218
Two systems monitored, new data acq. card	\$9,470	\$22,218
Three systems monitored, current data acq. card	\$9,486	\$22,218
Three systems monitored, new data acq. card	\$9,101	\$22,218

The largest single cost during the lifetime of the sewage system without a NILM installed is the replacement cost of the vacuum pumps. Supposing that the pumps were to last the lifetime of the ship and would not need to be replaced, the lifetime costs without a NILM drops to \$9,505. By comparison to the numbers in Table 6-4, it can be seen that there are still cost savings associated with installing a NILM if at least two systems are monitored using the new data acquisition card or at least three systems are monitored using the current data acquisition card.

## 6.4 Conclusions

The above calculations show that there are potential cost benefits to using a NILM on the sewage system. Although the assumptions made are not guaranteed to be accurate for an actual sewage system over the next thirty-eight years, the results are promising. A tool, such as a NILM, that can avert wasting energy has great potential in saving money throughout the lifetime of the ship.

## **7 Future Work and Conclusion**

### **7.1 System Modeling**

The Simulink and Matlab model introduced in this thesis performs well to simulate the operation of the sewage system onboard USCGC Seneca. The methodology used to create the model should be applicable to any system, thus the current model should be easily modifiable in order to replicate results from similar systems. Possible candidates for other systems to study are sewage systems onboard other Coast Guard Cutters, potable water systems, pressurized air systems and any other pressurized systems that use electrically driven pumps to “recharge” the system.

### **7.2 Diagnostic Indicator Testing**

The method developed and outputs provided by the diagnostic scheme introduced need in-service testing. The reliability and robustness of the diagnostics are based on the data collected over the last few years and analyzed after-the-fact. The diagnostic methods need to be applied and tested in real-time situations. A diagnostics module needs to be programmed immediately and installed onboard Seneca.

Diagnostic indicators from other components or systems not specifically discussed above could be useful if included in the diagnostics module. One such indicator would be the discharge pumps, introduced in section 2.3 above, that are also monitored by the current NILM installation onboard Seneca. The pumps operate approximately 1-2 times per day while underway. With increased system usage rates the discharge pumps would operate more often while a leak in the system would have no effect on the discharge pumps operating schedule. For any system onboard the ship, there are likely diagnostic indicators that can be deemed from interactions with other systems. Although it is an objective to keep the NILM as autonomous as possible, more input from other sources, including other NILMs, could be used if required to increase the robustness and reliability of some diagnostics indicators.

An extended period of testing, both at sea and inport, with feedback from the crew who monitor the system and receive the system status reports from the NILM, is required to conclude the theory-to-practice evolution of the NILM system.

Future potential NILM applications can possibly replace numerous sensors. NILM installations on shipboard systems that already contain pressure sensors, excessive-run indicators, etc. are necessary to further the NILM project. Testing NILM against legacy system monitors has the potential to show the usefulness of a single-point diagnostic tool.

### 7.3 Cost Considerations

Further development and testing of less expensive NILM components should continue. The cost savings associated with replacing the current data acquisition card with a less expensive design are great. A more economical current transducer design would also reduce the lifetime costs associated with a NILM installation.

A more detailed cost analysis needs to be performed. Mass production and competition between suppliers was not considered for the analysis done in this thesis. More in-depth calculations should be performed in order to better predict the lifetime costs of an actual NILM installation.

### 7.4 NILM Equipment

A custom designed and built data system which integrates the data acquisition, hard drive, and processing capabilities into one unit would be ideal. Not only would an integrated system possibly ease costs, but it would also save space onboard the ship. The current setups with full-sized PCs that are not watertight or very shock resistant should be replaced with more durable and compact units. Future applications will require smaller units with equal capacity to perform data collection, storage and analysis.

Research of NILM applications monitoring multiple systems is needed immediately. A natural progression from this point would be to use the sewage system NILM to monitor another system as well. Multi-tasking applications are vital to the success of the NILM project.



## 7.5 Conclusion

The Non-Intrusive Load Monitor installed in shipboard systems shows great promise for future applicability. The single-point connections required for a NILM and the capability to monitor multiple systems are ideal for the shipboard environment. A NILM can provide inputs to other systems if needed and can provide indications of normal or abnormal system operations. The diagnostic capabilities of a NILM can rival those of legacy installed monitoring devices.

Used alone or in conjunction with other systems, the NILM will be an important shipboard tool in the near future.

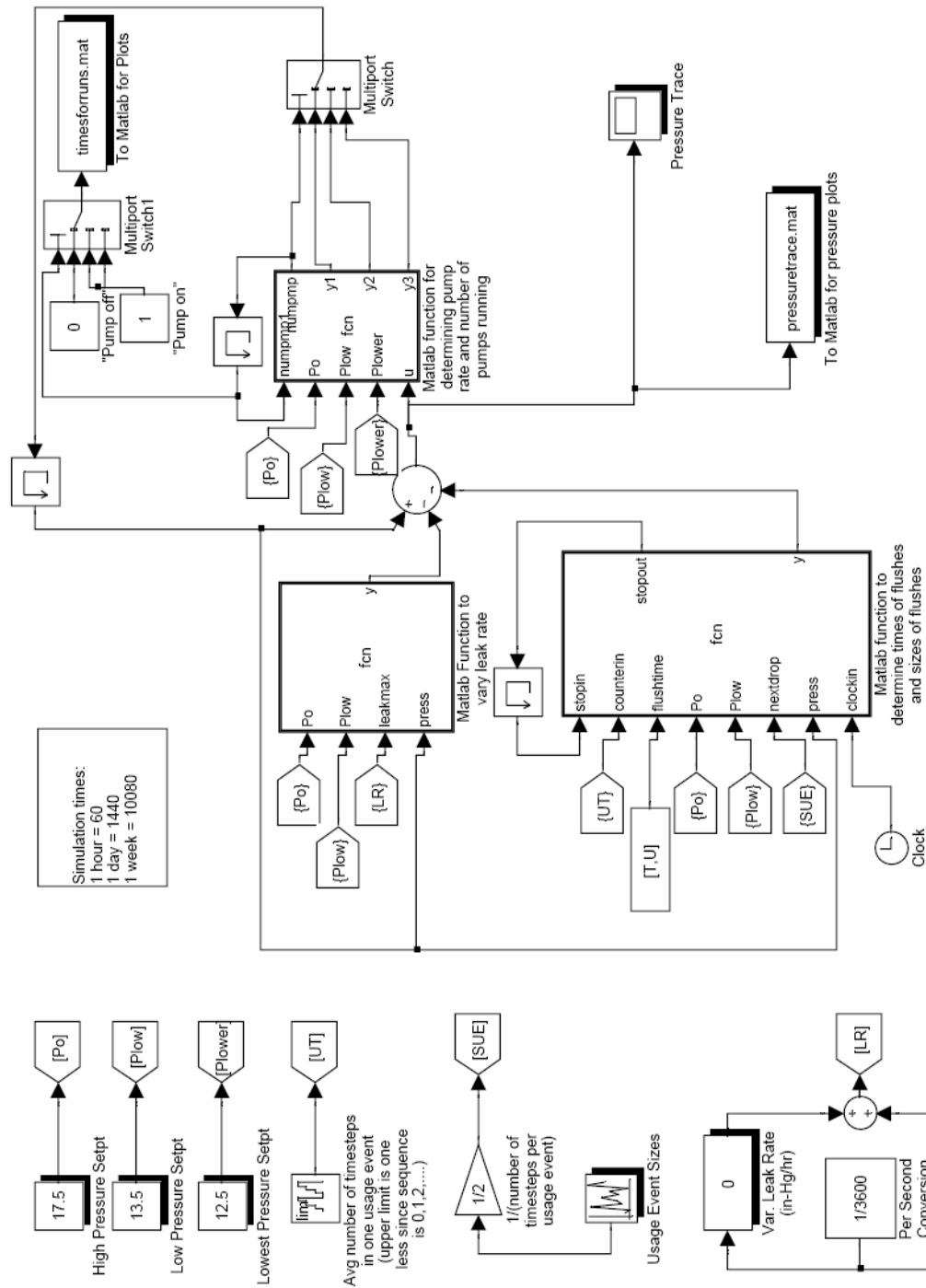
## List of References

- [1]. S.B. Leeb. "A Conjoint Pattern Recognition Approach to Non-Intrusive Load Monitoring," Ph.D. Dissertation, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 1993.
- [2]. S.R. Shaw, D. Evangelista, S.B. Leeb and C.R. Laughman. "Non-Intrusive Load Monitoring and Identification in an Automotive Environment," In Proceedings of ELECTRIMACS 1999, Lisbon, Portugal, pp. 199-204, September 1999.
- [3]. L.K. Norford and S.B. Leeb. "Nonintrusive Electrical Load Monitoring," Energy and Buildings, Vol. 24, pp. 51-64.
- [4]. S.B. Leeb, S.R. Shaw and J.L. Kirtley. "Transient event detection in spectral envelope estimates for nonintrusive load monitoring," IEEE Trans. Power Deliver, vol. 10, no. 3, pp. 1200-1210, July 1995.
- [5]. S.R. Shaw. "System Identification and Modeling for Nonintrusive Load Diagnostics," Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2000.
- [6]. J.S. Ramsey, Jr. "Shipboard Applications of Non-Intrusive Load Monitoring", Massachusetts Institute of Technology NSEE/S.M. EECS thesis, June 2004.
- [7]. T.W. DeNucci. "Diagnostic Indicators for Shipboard Systems using Non-Intrusive Load Monitoring," Massachusetts Institute of Technology S.M. NAME/S.M. ME Thesis, June 2005.
- [8]. W.C. Greene. "Evaluation of Non-Intrusive Monitoring for Condition Based Maintenance Applications on US Navy Propulsion Plants," Massachusetts Institute of Technology S.M. NAME/S.M. ME Thesis, June 2005.
- [9]. S.R. Shaw and C.R. Laughman. "A Kalman-Filter Spectral Envelope Preprocessor," submitted for publication in IEEE Trans. Instrum. Meas., February 26, 2004.
- [10]. R.W. Cox, S.B. Leeb, S.R. Shaw and L.K. Norford. "Transient event detection for non-intrusive load monitoring and demand side management using voltage distortion," in Proc. 21st Applied Power Electronics Conference (APEC), March 2006.
- [11]. USCGC Seneca website, "About Seneca," <http://www.uscg.mil/d1/units/cgcseneca/about%20seneca.htm>
- [12]. D.P. Bertsekas and J.N. Tsitsiklis, Introduction to Probability, Belmont, MA: Athena Scientific, 2002.

- [13]. A. Papoulis. Probability, Random Variables, and Stochastic Processes, New York, NY: McGraw-Hill, 1965.
- [14]. J. Paris. "A Framework for Non-Intrusive Load Monitoring and Diagnostics," Massachusetts Institute of Technology S.M. EECS Thesis, February 2006.
- [15]. Envirovac Inc. Instruction Manual; Installation, Operation and Maintenance with Parts List; Envirovac Vacuum Sewage Collection System; United States Coast Guard; U.S.C.G. 270' B Class WMEC. April 1983.
- [16]. G. Cowen. Statistical Data Analysis. New York, NY: Oxford University Press, 1998.
- [17]. LCDR Clyburn, USCG, Phone conversation, 19 April 2006.
- [18]. LCDR Mike Obar, USCG, Email correspondence, 06 April 2006.
- [19]. Advantech website. <http://www.advantech.com/>
- [20]. NIST/SEMATECH e-Handbook of Statistical Methods, <http://www.itl.nist.gov/div898/handbook/eda/section3/eda35f.htm>, 12 April 2006.
- [21]. N.K. Bose and C.R. Rao. Handbook of Statistics 10: Signal Processing and Its Applications. Amsterdam, The Netherlands: North-Holland, 1993.
- [22]. K.V. Bury. Statistical Models in Applied Science. New York, NY: John Wiley & Sons, 1975.
- [23]. P.R. Bevington and D.K. Robinson. Data Reduction and Error Analysis for the Physical Sciences. Second Edition, McGraw-Hill, 1992.
- [24]. J.A. Nachlas. Reliability Engineering: Probabilistic Models and Maintenance Methods. CRC Press, 2005.

Page Intentionally Left Blank

## Appendix A: Simulink Model and Matlab Embedded Code



## Embedded Functions:

1. Matlab function to vary leak rate.

```
function y = fcn(Po,leakmax,press)

% From the linear relation based on  $dP/dt = -k \cdot P$ 
leakmod = leakmax/Po*press;

% Output
y = leakmod;
```

2. Matlab function to determine times of flushes and sizes of flushes.

```
function [stopout,y] = fcn(stopin,counterin,flushtime,Po,Plow,nextdrop,press,clockin)
% This block places a flush into the system
% The code compares the sim time with the times generated in the
% sewage_model_rev1_prep.m file and inserts a flush where dictated by the T and U
% vectors.

counterin1 = double(counterin);

% To account for predominance of two flush sizes. If 100*clock in is even, the first
% factor is used and the other is used if it is odd.
if rem(ceil(100*clockin),2) == 0
    nextdropfactor = 1;
else
    nextdropfactor = 0.75;
end

% To account for flushes occurring over multiple time steps.
if flushtime == 1
    y=nextdrop*nextdropfactor;
    stopout=counterin1;

elseif stopin ~= -1
    y=nextdrop*nextdropfactor;
    stopout=stopin;
    if counterin1 == stopin
        y=0;
        stopout=-1;
    end
else
    y=0;
    stopout=-1;
end
```

3. Matlab function for determining pump rate and number of pumps running.

```
function [numpmp,y1,y2,y3] = fcn(numpmp1,Po,Plow,Plower,u)
% This block determines the output of the pumps, if any

% Definition of pumping rates
pump1rate = 4.5; %in-Hg/min
pump2rate = 4.7; %in-Hg/min

% These are the pumping rates
y1 = u;
y2 = u + pump1rate/60;
y3 = u + (pump1rate+pump2rate)/60;

% Check to see if the pumps should be turned off or how many should be on
if u >= Po
    numpmp = 0;
elseif u <= Plow && u > Plower
    if numpmp1 == 0
        numpmp = 1;
    else
        numpmp = numpmp1;
    end
elseif u <= Plower
    if numpmp1 ~= 2
        numpmp = 2;
    else
        numpmp = numpmp1;
    end
else
    numpmp = numpmp1;
end
```

Page Intentionally Left Blank



## Appendix B: Matlab Code

### 1. Matlab “prep” routine for Simulink simulation.

```
%This m-file is the prep routine for setting up to run the
%sewage_model_rev1 Simulink model.

clear;
% First get the required inputs
T=input('What is the simulation time you intend to run (in minutes)?');
lambda_w=input('What is the lambda value for the workday (in flushes/hour)?');
lambda_e=input('What is the lambda value for the evening (in flushes/hour)?');
lambda_n=input('What is the lambda value for the night (in flushes/hour)?');
perc_var=input('What is the percent variation for the lambda values (in %)?');
perc_var=perc_var/100;
filename = 'test';%input('What is the file name to save this run under?', 's');

% Now set it up to run for a third of the time with each lambda
flush_times=[];
t = 0;
rand('state',sum(100*clock));%97531
while t <= T/3
    t = t - 60*(log(rand) / lambda_w/(1+perc_var*rand(1))/(1-perc_var*rand(1)));
    flush_times=[flush_times,[t;1]];
end
while t > T/3 && t <= 2*T/3
    t = t - 60*(log(rand) / lambda_e/(1+perc_var*rand(1))/(1-perc_var*rand(1)));
    flush_times=[flush_times,[t;1]];
end
while t > 2*T/3 && t <= T
    t = t - 60*(log(rand) / lambda_n/(1+perc_var*rand(1))/(1-perc_var*rand(1)));
    flush_times=[flush_times,[t;1]];
end
save('flush_times'); %for use if want to compare different timesteps

% Now account for any errors that will occur if the flush times are too
% close together
timestep= 1/60;
j=2;
count=0;
while j <= length(flush_times)-1
    if flush_times(1,j)-flush_times(1,j-1) < timestep;
        flush_times(1,j) = flush_times(1,j)+timestep;
        count = count + 1;
    end
    if flush_times(1,j+1)-flush_times(1,j) < 0
        flush_times(1,j+1)=flush_times(1,j+1) + timestep;
    end
    j=j+1;
end
number_of_times_moved=count

%Now put this in a format that Simulink can understand and use
i=2;
times=[];
```

```

while i <= 2*length(flush_times)
    times(:,i-1) = flush_times(:,i/2);
    times(:,i) = [(flush_times(1,i/2)+timestep);0];
    i=i+2;
end

```

```

T = times(1,:);
U = times(2,:);

```

```

total_number_of_flushes = length(times)/2 - 1

```

```

% Now run the simulation and plot the results
sim('sewage_model_rev3');
sewage_model_post1

```

2. Matlab “post” routine for Simulink simulation

```

% This will read "timesforruns" and create the histogram
load('timesforruns.mat');

```

```

timeson = [];
timesoff = [];
timediff = [];
lnth = length(timesforruns(1,:));

```

```

% First find where the 1's and 0's change. 1 means pump on. 0 means off.

```

```

for i = 2:lnth
    if timesforruns(2,i-1) == 0 && timesforruns(2,i) == 1
        timeson = [timeson,timesforruns(1,i)];
    elseif timesforruns(2,i-1) == 1 && timesforruns(2,i) == 0
        timesoff = [timesoff, timesforruns(1,i)];
    end
end

```

```

if length(timeson) == length(timesoff)
    timediff = [timeson(1) timeson(2:end)-timesoff(1:end-1)];
else
    timediff = [timeson(1) timeson(2:end)-timesoff(1:end)];
end

```

```

%First, pullout the outliers and count them

```

```

j=1;
outlier_count = 0;
ln_timediff = length(timediff);
while j <= ln_timediff
    if timediff(j) >= 25
        timediff(j) = [];
        outlier_count = outlier_count+1;
        ln_timediff = length(timediff);
    else
        j=j+1;
    end
end
outlier_count

```

```

% Now do the histogram
[N,X] = hist(timediff,[0.25:0.5:25]);

```

```

N2 = medfilt1(N,7);
figure(1); clf;
bar(X,N);
hold on;
grid on;
xlabel('Time (min)');
ylabel('Counts');
set(gca,'FontName','Times','FontSize',14);
xl = get(gca,'XLabel');
yl = get(gca,'YLabel');
set(xl,'FontSize',18,'FontName','Times');
set(yl,'FontSize',18,'FontName','Times');
%plot(X,N2,'r','Linewidth',2);
hold off;

```

```

total_number_of_runs = sum(N)

```

```

gfit = gamfit(timediff);
k = gfit(1)
lambda = 1/gfit(2)

```